Constant scalar curvature metrics on toric surfaces

S. K. Donaldson

May 2, 2008

Contents

1 Introduction

This paper continues the series [3], [4], [5] in which we study the scalar curvature of Kahler metrics on toric varieties and relations with the analysis of convex functions on polytopes in Euclidean space. The main result of the present paper is an existence theorem for metrics of constant scalar curvature on toric surfaces, confirming a conjecture in [3].

We begin by recalling the background briefly: more details can be found in the references above. Let P be a bounded open polytope in $\mathbf{R}^n$ and let σ be a measure on the boundary of P which is a multiple of the standard Lebesgue measure on each face. Let A be a smooth function on the closure $\overline{P}$ of P and consider the linear functional $L_{A,\sigma}$ on the continuous functions on $\overline{P}$ given by

$$L_{A,\sigma} f = \int_{\partial P} f \ d\sigma - \int_P Af \ d\mu,$$

where $d\mu$ is ordinary Lebesgue measure on $\mathbf{R}^n$. We define a nonlinear functional on a suitable class of *convex* functions u on P by

$$\mathcal{M}(u) = -\int_P \log\det(u_{ij}) \ d\mu + L_{A,\sigma} u,$$

where (u_{ij}) denotes the Hessian matrix of second derivatives of u.

The Euler-Lagrange equation associated to $\mathcal{M}$ is the fourth order PDE found by Abreu:

$$\sum_{i,j} \frac{\partial^2 u^{ij}}{\partial x_i \partial x_j} = -A, \tag{1}$$

where (u^{ij}) is the matrix inverse of (u_{ij}). (We often use the notation u^{ij}_{ij} for the left-hand side of (1).) A solution of this equation, with appropriate boundary behaviour, is a critical point, in fact the minimiser, of $\mathcal{M}$. More precisely, we require u to satisfy *Guillemin boundary conditions*, which depend on the measure σ. We refer to the previous papers cited for the details of these boundary conditions and just recall here the standard model for the behaviour at the boundary. This is the function

$$u(x_1, \dots, x_n) = \sum x_i \log x_i,$$

on the convex subset $\{x_i > 0\}$ in $\mathbf{R}^n$. For appropriate "Delzant" pairs (P, σ) a function u with this boundary behaviour defines a Kahler metric on a corresponding toric manifold X_P. This comes with a map $\pi : X_P \to P$ and the scalar curvature of the metric is $A \circ \pi$. Thus when A is constant the Kahler metric has constant scalar curvature. When A is an affine-linear function the Kahler metric is "extremal".

The Guillemin boundary conditions imply that

$$L_{A,\sigma}(f) = \int_P \sum f_{ij} u^{ij} d\mu, \tag{2}$$

for smooth test functions f. Thus L_A vanishes on affine-linear functions. This just says that $(P, Ad\mu)$ and $(\partial P, d\sigma)$ have the same mass and moments. More interestingly, equation (2) tells us that $L_{A,\sigma}(f) \geq 0$ for smooth *convex* functions f, with strict equality if f is not affine-linear. This can be extended to more general convex functions f. In [3] we conjectured that this necessary condition for the existence of a solution u, is also sufficient. The present paper completes the proof of this in the case when the dimension n is 2 and the function A is a constant. Thus we have

Theorem 1 *Suppose $P \subset \mathbf{R}^2$ is a polygon and σ is a measure on ∂P, as above, with the property that the mass and moments of $(\partial P, \sigma)$ and $(P, Ad\mu)$ are equal for some constant A. Then either there is a solution to (1), satisfying Guillemin boundary conditions, or there is convex function f, not affine-linear, with $L_{A,\sigma}(f) \leq 0$*

A corollary in the framework of complex geometry of this is

Corollary 1 *If a polarised complex toric surface with zero Futaki invariant is K-stable it admits a constant scalar curvature Kahler metric.*

We refer to [3] for the terminology and further details. (The converse is also true and is proved by Zhou and Zhu in [9].)

Two remarks are in order here. First, in the two-dimensional case the "positivity" condition for the functional L_A can be made more explicit. If λ is an affine-linear function on $\mathbf{R}^2$ define

$$\lambda^+(\underline{x}) = \max(0, \lambda(\underline{x})).$$

Then it follows from the arguments of [3] that the second alternative in Theorem is equivalent to the condition that there is some function of the form λ^+, not identically zero on P, with $L_{A,\sigma}(\lambda^+) \leq 0$. Thus to check the existence of a solution to (1) we only need to check the positivity of $L_{A,\sigma}$ on a 2-parameter family of functions of the form λ^+, which would be easy to do with a computer. Second, it was shown in [3] that the positivity condition is equivalent to a more useful, quantitative, statement. Fix a base point p_0 and say a convex function f is *normalised at* p_0 if p_0 is the minimiser of f and $f(p_0) = 0$. Then the positivity property holds if and only if there is some $C = C_{A,\sigma,p_0} > 0$ such that

$$\int_{\partial P} f d\sigma \leq C \; L_{A,\sigma} f \tag{3}$$

for all normalised convex functions f.

The main result of [5] asserts that Theorem 1 follows if one can establish a certain *a priori* estimate on solutions u. Suppose that (3) holds and u is a solution, normalised at p_0 Then taking $f = u$ in (2) we obtain

$$\int_{\partial P} u \; d\sigma \leq C \; L_{A,\sigma} u = 2C\text{Area } (P). \tag{4}$$

This immediately gives *a priori* C^1 bounds on the restriction of u to compact subsets of P, which can be applied to bound all derivatives in the interior, as in [4]. In [5] we showed that the estimates could be extended up to the boundary provided that u satisfies a condition, called there an "M-condition". We recall the definition. Let p, q be points in P and write $q = p + d\ \nu$ where ν is a unit vector. Write $V(p, q)$ for the difference in the derivative of u in the ν direction evaluated at q and p. Then u satisfies an M condition if for any pair p, q in P such that the points $p - d\ \nu$ and $q + d\ \nu$ also lie in P we have $V(p, q) \leq M$.

Roughly speaking, the point of the present paper is to show that the solutions of our equation satisfy an *a priori* M-condition. But the detailed strategy is considerably more complicated and we outline it now.

1. In Section 2, we obtain an *a priori* L^∞ bound on the solutions. (Notice that an M-condition implies such a bound, by an easy argument.)

2. In Section 3, we work with points p in the neighborhood of an edge of the polygon but away from the vertices. We obtain an *a priori* bound on a quantity $D(p)$, which is essentially equivalent to the M-condition for pairs of points p, q not close to the the vertices. This is the core of the paper: we use a "blow-up" argument hinging on the compactness properties of sets of bounded convex functions.

3. In Section 4 we use the arguments of [5] to obtain bounds on the curvature tensor of our solution, away from the vertices. This requires some subsidiary arguments to control the Riemannian distance function, also using the bound on the quantity $D(p)$ from Section 3.

4. In Section 5 we study the solutions in neighbourhoods of the vertices. First, using the maximum principle and our L^∞ bound from Section 2, we obtain a two-sided bound on the volume form $\det u_{ij}$. We then use arguments similar to those in Section 3 to obtain an *a priori* bound on a quantity $E(t)$ which is related to the M-condition near the vertices.

Many steps in this programme apply equally well to the general problem, with other functions A. The main obstacle in making this extension comes in the arguments of Section 3. Here we use a special estimate, for the case of constant A, proved in [4]. If u is any smooth convex function we can define a vector field V by

$$V^i = -\sum_j \frac{\partial u^{ij}}{\partial x_j}. \tag{5}$$

In [5], Theorem 2 we showed that, when the dimension is 2 and the function A is constant, there is an *a priori* bound on the Euclidean norm of V,

$$|V|_{\mathrm{Euc}} \leq C. \tag{6}$$

Apart from the use of this estimate in Section 3, the only other restriction on A comes when, in two places, we use the fact that A is positive (in arguments

involving the maximum principle) Thus the main result in this paper extends to positive functions A if one assumes an *a priori* estimate (6). However the uses we make of (6) are to overcome some rather technical difficulties which do not appear to be central to the problem, so there are good grounds for hoping that the proofs may be extended to other functions A in time.

Let us say a little more here about the central argument in Section 3 of the paper. The M-condition is essentially a local C^0 bound on a convex function. Suppose u is a smooth convex function normalised at the origin. Then we can obviously choose a small scalar ϵ such that $u_\flat = cu$ satisfies a fixed C^0 bound on a fixed ball about the origin and we have $(u_\flat)^{ij}_{ij} = c^{-1} u^{ij}_{ij}$. Then if we have a sequence $u^{(\alpha)}$ such that $\left(u^{(\alpha)}\right)^{ij}_{ij}$ is suitably small we can make a sequence of such re-scalings to get a new sequence $u_\flat^{(\alpha)}$, bounded in C^0 and with $\left(u_\flat^{(\alpha)}\right)^{ij}_{ij}$ tending to zero. We can suppose the $u_\flat^{(\alpha)}$ converge in C^0 over the interior of the ball and the question is: what can we say about the limit? This is the basic idea used in Section 3, and also (more implicitly) in Section 5. The complexities of the arguments, and the restrictions on the results, are related to our lack of understanding of this basic local question—what are the possible C^0 limits of solutions to the equation (1)?

The main work of this paper finishes in Section 5. However, it is interesting to complement the existence proofs with an understanding of what goes wrong when the positivity hypothesis is violated. Section 6 is a supplement in which we discuss an explicit family of complete zero scalar curvature metrics, generalising the Taub-NUT metric, and explain that these can be expected to arise as blow-up limits of solutions. We also discuss briefly the connection with "collapsing" phenomena in Riemannian geometry.

2 The L^∞ estimate

2.1 An integral inequality

In this subsection we derive a geometric inequality for solutions of the equation (1). Consider the general case of a polytope $P \subset \mathbf{R}^n$ with boundary measure σ and a function A. We suppose that u is a solution to (1) satisfying Guillemin boundary conditions. Recall that this implies that for any smooth test function f on $\overline{P}$ we have

$$\int_{\partial P} f d\sigma = \int_P u^{ij} f_{ij} + A f d\mu.$$

Suppose that $\underline{u}$ is a weakly-convex smooth function on $\overline{P}$ and that $\underline{u} = u$ on an open subset $X \subset \overline{P}$. Take $f = u - \underline{u}$, so f_{ij} vanishes on X. Outside X we have

$$u^{ij} f_{ij} = u^{ij}(u_{ij} - \underline{u}_{ij}) \leq u^{ij} u_{ij} = n,$$

since $\underline{u}$ is convex. Thus

$$\int_{\partial P} u - \underline{u}\ d\sigma \leq n\mathrm{Vol}(P \setminus X) + \int_P A(u - \underline{u})\ d\mu. \tag{7}$$

Now start with an open set $X \subset \overline{P}$, with piecewise-smooth boundary, say. We define a convex function $\underline{u}_X$ by

$$\underline{u}_X = \max_{p \in X \cap P} \lambda_p,$$

where for each $p \in P$ we write λ_p for the affine-linear function defining the supporting hyperplane of the graph of u. Thus $\underline{u}_X = u$ on X, by convexity, and an alternative definition is that $\underline{u}_X$ is the least convex function which restricts to u on X. We claim that the inequality (6) holds, with $\underline{u} = \underline{u}_X$. This is immediate if $\underline{u}_X$ is smooth. In the general case, we introduce a small parameter ϵ and let $\underline{u}_{X,\epsilon}$ be the standard mollification of $\underline{u}_X$, using convolution with a bump-function supported in the ϵ- ball. (Notice that $\underline{u}_X$ is defined on all of $\mathbf{R}^n$.) Then $\underline{u}_{X,\epsilon}$ is again convex and is equal to u on the set $X_\epsilon \subset X$, defined by removing the ϵ-neighbourhood of the boundary of X. Then we can apply (6) to $\underline{u}_{X,\epsilon}$, so

$$\int_{\partial P} u - \underline{u}_{X,\epsilon} d\sigma \leq n\mathrm{Vol}(P \setminus X_\epsilon) + \int_P A(u - \underline{u}_{X,\epsilon})\ d\mu.$$

We take the limit as ϵ tends to 0. The functions $\underline{u}_{X,\epsilon}$ converge uniformly to $\underline{u}_X$ and the volume of X_ϵ tends to the volume of X by our regularity assumption on the boundary of X. So, in sum, we have derived the geometric inequality

$$\int_P u - \underline{u}_X\ d\mu \leq n\mathrm{Vol}(P \setminus X) + \int_P A(u - \underline{u}_X)\ d\mu. \tag{8}$$

(Notice that this inequality makes sense for arbitrary convex functions u. We can think of it as a partial "weak form" of the equation (1) , and the boundary conditions.)

2.2 L^∞ estimate: the main idea

In this subsection and the next we will apply the ideas above to derive an *a priori* bound for $\max \tilde{u}$, when $\tilde{u}$ is a normalised solution of (1), and the boundary conditions, assuming an L^1 bound on the restriction to the boundary. Of course this maximum is attained at one of the vertices, so we fix a vertex q and seek to bound $\tilde{u}(q)$. We can choose coordinates so that q is the origin and, near to q, the polygon P is the first quadrant $\{x_1 > 0, x_2 > 0\}$. We write $(2l_1, 0)$ and $(0, 2l_2)$ for the coordinates of the two vertices adjacent to q. We can arrange that the boundary measures on these two edges are standard. Recall that the L^1 bound on the boundary values of $\tilde{u}$ gives, by very elementary arguments, bounds on $\tilde{u}$ and its first derivative in the interior of each edge of the boundary. Let u be the unique function obtained by adding an affine-linear function to $\tilde{u}$ such that

- $\frac{\partial u}{\partial x_1} = -1$ at the midpoint $(l_1, 0)$ and $\frac{\partial u}{\partial x_2} = -1$ at $(0, l_2)$.
- The minimum value of u on P is 0.

All of these preliminaries are just to provide a convenient setting for the main arguments. Clearly, the bounds on the derivatives of u at the midpoints mean that it suffices to obtain an *a priori* bound on $u(0,0)$.

The goal of this section is to prove

Theorem 2 *There is an H depending only on $l_1, l_2, \|A\|_{L^\infty}$ and the integral of u over the boundary of P such that $u(0,0) \leq H$.*

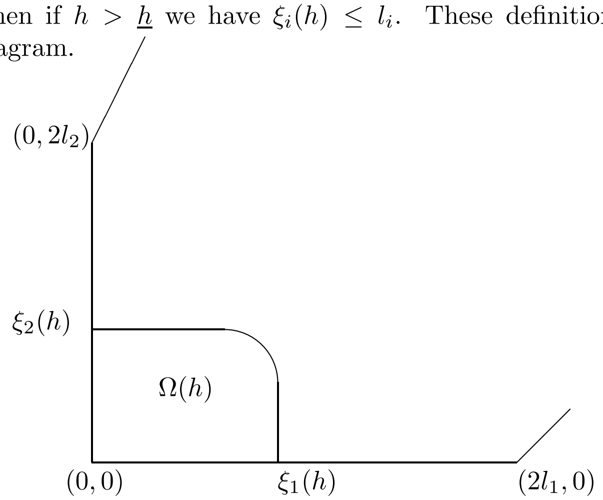
To prove the Theorem, we will apply our inequality (8) to a 1-parameter family of domains $X(h)$ in P. Define a function ϕ on P by

$$\phi = u - x_1 \frac{\partial u}{\partial x_1} - x_2 \frac{\partial u}{\partial x_2},$$

and set

$$X(h) = \{\underline{x} : \phi(\underline{x}) < h\}.$$

Thus $X(h)$ is the largest subset with the property that $\underline{u}_{X(h)}(0,0) \leq h$. As the parameter h increases the domain $X(h)$ grows and once $h \geq u(0,0)$ we have $X(h) = P$. We write $\Omega(h)$ for the complement $P \setminus X(h)$. We also write $\overline{h} = u(0,0)$ and work with values $h < \overline{h}$. Then the closure of $\Omega(h)$ meets the axes in a pair of line segments, from the origin to $(\xi_1(h), 0), (0, \xi_2(h))$ respectively, say. We also write $\underline{h}_1 = u(l_1, 0) + l_1$ and $\underline{h}_2 = u(0, l_2) + l_2$ and set $\underline{h} = \max(\underline{h}_1, \underline{h}_2)$. Then if $h > \underline{h}$ we have $\xi_i(h) \leq l_i$. These definitions are illustrated in the diagram.



Let $\tau_{1,h}(t)$ be the affine-linear function of one variable whose graph is the supporting hyperplane of the restriction of u to the x_1-edge at the point $\xi_1(h)$.

Thus, by definition, $\tau_{1,h}(0) = h$. It follows from the definition that the restriction of the function $\underline{u}_{X(h)}$ to the axis is supported on the interval $[0, \xi_1(h))$ on which it is equal to the affine-linear function $\tau_{1,h}$. Let

$$G_1(h) = \int_0^{\xi_1(h)} u(t,0) - \tau_{1,h}(t)\ dt,$$

and define $\xi_2(h)$ and $G_2(h)$ similarly. Then

$$\int_{\partial P} u - \underline{u}_{X(h)} d\sigma = G_1(h) + G_2(h).$$

We can now explain the main idea of our proof. To begin with let us suppose that for $h \geq \underline{h}$ the support of the function A does not meet Ω_h. Our basic inequality (8) becomes

$$G_1(h) + G_2(h) \leq \mathrm{Area}(\Omega_h).$$

Elementary calculus gives the identities

$$\frac{dG_i}{dh} = -\frac{1}{2}\xi_i(h). \tag{9}$$

In the standard model, where $u = x_1 \log x_1 + x_2 \log x_2 + \text{constant}$ say, it is easy to check that Ω_h is exactly the triangle with vertices $(0,0), (\xi_1(h), 0), (0, \xi_2(h))$, so in this case the area of Ω_h is $\xi_1\xi_2/2$. Suppose that, in our general situation, we were able to show that Ω_h is not too different from this triangle, in that we have an inequality

$$\mathrm{Area}(\Omega_h) \leq \kappa\ \xi_h \xi_2(h), \tag{10}$$

for some fixed κ. For example, if we knew that Ω_h is contained in the rectangle with vertices $(0,0), (\xi_1(h), 0), (0, \xi_2(h)), (\xi_1(h), \xi_2(h))$ we could take $\kappa = 1$. Under this supposition we have

$$G_1 + G_2 \leq \kappa\ \xi_1\xi_2 \leq \kappa \left(\frac{dG_1}{dh} + \frac{dG_2}{dh} \right)^2,$$

where we have used (9). So the positive, decreasing, function $\Gamma = G_1 + G_2$ satisfies the differential inequality

$$\frac{d\Gamma}{dh} \leq -\sqrt{\frac{\Gamma}{\kappa}}$$

in the interval $\underline{h} < h < \overline{h}$. This gives

$$\sqrt{\Gamma}(\underline{h}) - \sqrt{\Gamma}(h) \geq \frac{1}{\sqrt{\kappa}}(h - \underline{h}),$$

and thus, since Γ tends to 0 as h tends to $\overline{h}$,

$$\overline{h} - \underline{h} \leq \sqrt{\kappa \Gamma(\underline{h})}.$$

Since $\Gamma(\underline{h})$ is dominated by the integral of u over the boundary of P this gives the desired bound on $\underline{h} = u(0,0)$.

To turn this idea into a complete proof we need to overcome two difficulties. The first is to incorporate the term involving the function A in our basic inequality. This is relatively easy. The second, more fundamental, difficulty is that the author does not know how to obtain a universal inequality of the form (10) that we used above, although it seems very reasonable to expect this to be true. Thus the actual proof, which we give in the next section, is more complicated since it is based on a weaker assertion than (10) (Lemma 2 below).

2.3 The detailed proof

We begin with some elementary calculus associated to a convex function of one variable. This will be applied to the boundary values of our function u, but to simplify notation consider first a strictly convex, smooth, function $U(t)$ on an interval $(0, 2l)$ with $U'(l) = -1$. For $h \geq U(l)+l$ we define $\xi(h) \in (0,l)$ as above, i.e. so that the affine-linear function τ_h whose graph is tangent to the graph of U at ξ has $\tau_h(0) = h$. Let $D(h)$ be the point where the affine-linear function τ_h vanishes. In other words, the line joining the two points $(h,0)$ and $(0, D(h))$ is tangent to the graph of U at the point $(\xi(h), U(\xi(h))\ )$. Let $z(h) = D(h)/h$, so $z(h)^{-1} = -u'(\xi(h))$.

Lemma 1 *In this situation*

$$\xi(h) = \frac{D^2}{D - hD'}.$$

This is a calculus exercise for the reader.

Now return to our function u of two variables. We extend the notation above in the obvious way, so we have functions $D_i(h), z_i(h)$ defined for $\underline{h} \leq h \leq \overline{h}$. For these values of h, we let Δ_h be the triangle in the (x_1, x_2) plane with vertices $(0,0), (D_1(h), 0), (0, D_2(h))$.

Lemma 2 *For any $h \in (\underline{h}, \overline{h})$ we have $\Omega(h) \subset \Delta_h$.*

To see this, consider a point p in $\Omega(h)$. Let π be the affine-linear function defining the supporting hyperplane of u at the point p and write $h^* = \pi(0,0)$. The condition that p lies in Ω_h is the same as saying that $h^* \geq h$. Consider the restriction of π to the x_1-axis. By convexity we have $\pi(t,0) \geq u(t,0)$ for all t, in particular $h^* \leq u(0,0)$ and so $\xi(h^*)$ is defined. Then τ_{1,h^*} and the restriction of π are two affine-linear functions of one variable, equal at the origin. We must have $\pi(t,0) \geq \tau_{1,h^*}(t)$ for all $t > 0$, for otherwise $\pi(t,0) \leq \tau_{h^*}(t)$ for all $t > 0$, which is a contradiction when $t = \xi_1(h^*)$. Thus $\pi(D_1^*, 0) = 0$ for some $D_1^* < D_1(h)$. Similarly $\pi(0, D_2^*) = 0$ for some $D_2^* < D_2(h)$. Now $\pi(0,0) > h > 0$ so the region in P where $\pi > 0$ is the triangle Δ^* with vertices $(0,0), (D_1^*, 0), (0, D_2^*)$. At the original point p we have $u(p) = \pi(p)$. Since u

was normalised so that $u \geq 0$ we have $\pi(p) \geq 0$ and so p lies in Δ^*. But Δ^* is contained in Δ_h, since $D_i^* < D_i$, so p lies in Δ_h, as required.

Next we look at the term involving the function A. For $h \geq \underline{h}$ we write $f_h = u - \underline{u}_{X(h)}$. So f_h is a positive function, supported in the set Ω_h. We have to estimate the term

$$\int_{\Omega_h} A f_h,$$

appearing in the inequality (7). Set $\alpha = \|A\|_{L^\infty}$ so

$$\int_{\Omega_h} A f_h \leq \alpha J(h), \tag{11}$$

where

$$J(h) = \int_{\Omega_h} f_h. \tag{12}$$

Lemma 3 *With notation as above,*

$$\frac{dJ(h)}{dh} = -\frac{1}{3}\text{Area }(\Omega_h).$$

This is the two-dimensional analogue of the elementary identity (9). To prove it we work in polar coordinates, writing $u(r,\theta)$. Consider a ray through the origin, with fixed θ. The restriction of u to this ray is a convex function of r and there is a unique point $r = R(\theta)$ where $u - r\frac{\partial u}{\partial r} = h$. From the definitions, this is a point on the boundary $\partial\Omega_h$ and the restriction of the function $\underline{u}_{X(h)}$ to the intersection of Ω_h and this ray is

$$\underline{u}_{X(h)}(r,\theta) = h + \frac{r}{R}(u(R,\theta) - h).$$

It follows that

$$\frac{\partial}{\partial h}\underline{u}_{X(h)}(r,\theta) = 1 - \frac{r}{R}.$$

Thus

$$\frac{d}{dh}J(h) = \int_{\theta=0}^{\theta=\pi/4} \int_{r=0}^{r=R(\theta)} (1 - \frac{r}{R}) \; r dr d\theta.$$

Performing the r integral this is

$$\frac{d}{dh}J = \frac{1}{6}\int_{\theta=0}^{\theta=\pi/4} R(\theta)^2 d\theta,$$

while the area of Ω_h is given by the usual formula

$$\text{Area }(\Omega_h) = \frac{1}{2}\int R(\theta)^2 d\theta.$$

Combining the two lemmas above, we get

Corollary 2

$$J(h) \leq \frac{1}{6}\int_{h}^{\overline{h}} D_1(h)D_2(h)\ dh.$$

This follows immediately from the co-area formula and the facts that the area of Δ_h is $D_1D_2/2$ and that $J(h) \to 0$ as $h \to \underline{h}$.

Now define $I_1(h)$, for $\underline{h} \leq h < \overline{h}$ by

$$I_1(h) = \int_{h}^{\overline{h}} D_1(h)^2\ dh, \tag{13}$$

and define $\lambda_1(h)$ by the equation

$$G_1(h) = \frac{\lambda_1(h)}{2}D_1(h)^2 + \frac{\alpha}{12}I(h). \tag{14}$$

(The reason for the choice of factor $\alpha/12$ will appear shortly.)

Define I_2 and λ_2 similarly.

Lemma 4 *$\lambda_i(h) \to 0$ as $h \to \overline{h}$*

This is straightforward to check, using the known behaviour of u at the origin. We omit the details.

Now we can proceed to the core of the proof, which has two parts. The first is stated in

Proposition 1 *Suppose $\underline{h} \leq h \leq \overline{h}$ and $\lambda_1(h), \lambda_2(h)$ are both positive. Then*

$$\lambda_1(h)\lambda_2(h) \leq 1.$$

Using Lemma 2, (11) and Corollary 2, our basic inequality (8) gives, for any $h \in (\underline{h}, \overline{h})$,

$$G_1(h) + G_2(h) \leq \frac{1}{2}D_1(h)D_2(h) + \frac{\alpha}{6}\int_{h}^{\overline{h}} D_1(s)D_2(s)\ ds.$$

Using the inequality $D_1D_2 \leq \frac{D_1^2+D_2^2}{2}$ and the definition of $I_i(h)$ we get

$$G_1(h) + G_2(h) \leq \frac{1}{2}D_1(h)D_2(h) + \frac{\alpha}{12}(I_1(h) + I_2(h)). \tag{15}$$

So we have, from the equations defining λ_i,

$$\frac{\lambda_1}{2}D_1^2 + \frac{\lambda_2}{2}D_2^2 \leq D_1D_2,$$

for each $h \in (\underline{h}, \overline{h})$. In other words

$$\frac{1}{2}\left(\lambda_1\frac{D_1}{D_2} + \lambda_2\frac{D_2}{D_1}\right) \leq 1.$$

From the arithmetic-geometric mean inequality we see that if $\lambda_1(h), \lambda_2(h)$ are both positive then $\lambda_1\lambda_2 \leq 1$.

For the second part, we derive differential equations involving the functions $\lambda_i(h), z_i(h)$. For clarity we suppress the suffix i temporarily, and denote derivatives with respect to h by a prime symbol. Write $c = \alpha/12$ and recall that $G = \frac{\lambda}{2}D^2 + cI$, where $\frac{dI}{dh} = -D^2$. Thus, using (9),

$$-\frac{\xi}{2} = G' = \lambda DD' + \frac{1}{2}\lambda' D^2 - cD^2.$$

By Lemma 1 this gives,

$$\frac{D^2}{2(hD' - D)} = \lambda DD' + \frac{1}{2}\lambda' D^2 - cD^2.$$

Since $D = zh$, we have

$$hD' - D = h^2 z' < 0,$$

and our equation becomes

$$\frac{z^2}{2z'} = \lambda zh(hz' + z) + \frac{\lambda'}{2}z^2h^2 - cz^2h^2.$$

This leads to

$$\frac{z}{2h^2} = z'\left(\lambda z' + z(\frac{\lambda}{h} - c + \frac{\lambda'}{2})\right).$$

Recall that $z > 0$ and $z' < 0$. Suppose K is any fixed positive number. For any $A, B > 0$ we have $K\sqrt{AB} \leq \frac{1}{2}(K^2A + B)$. We apply this to the right hand side of (15), with $A = -z'$ and $B = -(\lambda z' + z(\lambda/h - c + \lambda'/2)$. We deduce that

$$z'(K^2 + \lambda) + z(\frac{\lambda}{h} - c + \frac{\lambda'}{2}) \leq -\sqrt{2}K\frac{\sqrt{z}}{h}. \tag{16}$$

Proposition 2 *Suppose $z(h), \lambda(h)$ are functions defined on an interval $(h_0, \overline{h})$ with the following properties*

1. *$z(h) > 0$ and $z'(h) < 0$ for all h.*
2. *z, λ satisfy the differential inequality (16) above, for some $c > 0$ and all $K > 0$.*
3. *For some $C > 0$, and all h, we have*

$$z \leq C/h^2.$$

4. *$z(h)$ and $\lambda(h)$ tend to 0 as $h \to \overline{h}$.*

Write $b = 2\sqrt{2}-1$. If we fix any $K > 2c\sqrt{C}$ and if we set $h_1 = \max(h_0, \frac{3K\sqrt{C}}{b})$, then we have $K^2 + \lambda(h) > 0$ for all $h \geq h_1$ and

$$\int_{h_1}^{\underline{h}} \frac{1}{(K^2+\lambda)^{3/4}} \frac{dh}{h} \leq \frac{12}{bK}\ z(h_1)^{1/2}(K^2+\lambda(h_1))^{1/4}.$$

We fix K as stated. Multiplying the inequality (16) by $2z$, we have

$$2zz'(K^2+\lambda) + \lambda' z^2 + 2z^2\left(\frac{\lambda}{h} - c\right) \leq -\frac{2\sqrt{2}K}{h} z^{3/2}.$$

This is

$$\frac{d}{dh}\left(z^2(K^2+\lambda)\right) \leq -\frac{2\sqrt{2}K}{h} z^{3/2} + 2z^2\left(c - \frac{\lambda}{h}\right).$$

Now the inequalities $z \leq Ch^{-2}$ and $K > 2c\sqrt{C}$ imply that $2z^2 c < \frac{K}{h} z^{3/2}$ so we have

$$\frac{d}{dh}\left(z^2(K^2+\lambda)\right) \leq -\frac{bK}{h} z^{3/2} - \frac{2\lambda}{h} z^2. \tag{17}$$

Write $F = K^2 + \lambda$. We want to show that $F(h)$ is positive for $h \geq h_1$, so we suppose that $F(h_2) < 0$ for some $h_2 > h_1$ and seek a contradiction. Since $F \to K^2 > 0$ as $h \to \overline{h}$ there is an $h_3 \geq h_2$ with $F(h_3) = 0$ and $F'(h_3) \geq 0$. This means that

$$\frac{d}{dh}(z^2 F) = z^2 F' + (z^2)' F$$

is positive when $h = h_3$. So, with $\lambda = \lambda(h_3), z = z(h_3)$ we have

$$2\frac{\lambda}{h} z^2 \leq -\frac{bK}{h} z^{3/2}.$$

But $\lambda = -K^2$ (since $F = 0$) and we get

$$K z^{1/2} \geq \frac{b}{2}.$$

Since $z \leq C/h_3^2$, we have

$$\frac{\sqrt{C}}{h_3} \geq \frac{b}{2K}.$$

But this contradicts the assumption that $h_1 \geq 3K\sqrt{C}/b$, since $h_3 \geq h_1$. So we have established that $F(h) > 0$ for $h > h_1$.

Now if $h > h_1$ we have

$$-\frac{\lambda}{h} \leq \frac{K^2}{h} \leq \frac{K^2}{h_1}$$

and so

$$-2z^2\frac{\lambda}{h} \leq \frac{2\sqrt{C}K^2}{h_1}\ \frac{z^{3/2}}{h}.$$

So we obtain from (17) that

$$\frac{d}{dh}(z^2F) \leq (\frac{2K^2\sqrt{C}}{h_1} - bK)\frac{z^{3/2}}{h}.$$

By the choice of h_1, this gives,

$$\frac{d}{dh}(z^2F) \leq -\frac{bK}{3}\frac{z^{3/2}}{h}.$$

Write $w = z^2F$ so the above inequality is

$$\frac{dw}{dh} \leq -\frac{bK}{3h}\frac{w^{3/4}}{F^{3/4}}.$$

That is

$$\frac{d}{dh}w^{1/4} \leq -\frac{bK}{12}F^{-3/4}\frac{1}{h}.$$

We know that $w(h)$ tends to zero at $\underline{h}$ so we can integrate this with respect to h to obtain the inequality stated in Proposition 2.

Propositions 1 and 2 are the essential parts of the proof of Theorem 2, and it remains now to put together the various components. Of course we want to apply Proposition 2 to the functions λ_i, z_i associated to the two edges, with $h_0 \geq \underline{h}$. We have to show that there is a bound $z_i \leq Ch^{-2}$. Consider the quantity $z_1h^2/2 = hD_1(h)/2$. This is the integral of the affine-linear function $\tau_{1,h}$ from 0 to $D_1(h)$. The value $u(l_1, 0)$ is controlled by the integral of u over the boundary so there is no loss in supposing that $h_0 \geq 2u(l_1, 0)$. This implies that, for $h \geq h_1$ we have $D_1(h) \leq 2l_1$ and, since $\tau_{1,h}(t) \leq u(t, 0)$ for t in the interval $[0, D_1(h)]$, we get

$$\frac{1}{2}hD_1(h) \leq \int_{\partial P} u.$$

This gives the desired bound $z_1 \leq Ch^{-2}$, and similarly of course for z_2. Thus we may apply Proposition 2, with a suitable fixed K determined by C and $c = \|A\|_{L^\infty}/12$. Now consider the functions $w_i(h) = z_i^2(K^2 + \lambda_i(h))$. By construction $z_i(h) \leq 1$ for $h \geq \underline{h}$ so $w_i \leq K^2 + z_i^2\lambda_i = K^2 + \lambda_i D_i^2/h^2$. By the definition of λ_i we have

$$\lambda_i D_i^2 = 2(G_i - cI_i) \leq 2G_i.$$

Obviously the functions G_i are bounded by the integral of u over ∂P. So we obtain

$$w_i \leq K^2 + \frac{2}{h^2}\int_{\partial P} u.$$

This gives an upper bound on $w_i(h_1)$, since h_1 is determined by C and c. In sum, Proposition 2 tells us that there are h_1, L, K, all determined by the integral of u over the boundary, such that $K^2 + \lambda_i(h) > 0$ if $h > h_1$ and

$$\int_{h_1}^{\overline{h}} \frac{1}{(K^2+\lambda_i)^{3/4}}\frac{dh}{h} \leq L.$$

Change variable by writing $h = e^t$ and let $\overline{h} = e^{\overline{t}}, h_1 = e^{t_1}$. Then we have

$$\int_{t_1}^{\overline{t}} \frac{1}{(K^2 + \lambda_i)^{3/4}} \, dt \leq L.$$

So the measure of the set in $[t_1, \overline{t}]$ where $\lambda_i \leq 1$ is at most $L(K^2 + 1)^{3/4}$. Thus if $\overline{t}$ were bigger than $t_1 + 2L(K^2 + 1)^{3/4}$ there would have to be a point where $\lambda_1 > 1$ and $\lambda_2 > 1$. But this would contradict Proposition 1. So we conclude that $\overline{t} \leq t_1 + 2L(K^2 + 1)^{3/4}$ or in other words

$$\overline{h} \leq h_1 \exp(2L(K^2 + 1)^{3/4}),$$

and we have proved Theorem 2.

3 Edges

3.1 Preliminaries

We now come to the central topic of this paper. Consider a symplectic potential u on a polygon P and let E be an edge of P. Choose an outward-pointing vector ν transverse to E—say the Euclidean normal. Suppose that p is a point of P such that the ray $\{p + t\nu : t > 0\}$ meets the edge E in a point $q = p + s\nu$, for $s = s(p)$. Let λ_p be the affine-linear function defining the supporting hyperplane to u at p; so the difference $u - \lambda_p$ vanishes to first order at p. Then we define

$$D(p) = \frac{u(q) - \lambda_p(q)}{s(p)}. \tag{18}$$

The goal of this section is to obtain an *a priori* bound on $D(p)$, under mild hypotheses. It is easy and elementary to go from this to an "M-condition" formulation, as we will explain in Section 4. When we want to indicate the dependence on the function u we write $D(u; p)$.

We will want to have this *a priori* bound in the context of the continuity method of [5]; when we have a sequence $(P^{(\alpha)}, A^{(\alpha)}, \sigma^{(\alpha)})$ of data sets and convex functions $u^{(\alpha)}$. In this context the data sets will converge in the obvious sense as $\alpha \to \infty$. To simplify notation we will often omit the index α, and just write P, A, σ, where it is clear that the quantities involved (for example the diameter of $P^{(\alpha)}$) satisfy a uniform bound in the sequence. We suppose that $u^{(\alpha)}$ is normalised (in the sense of Section 1) at the centre of mass of $P^{(\alpha)}$. The main result we prove is

Theorem 3 *Suppose that the data sets $(P^{(\alpha)}, A^{(\alpha)}, \sigma^{(\alpha)})$ converge as $\alpha \to \infty$ and that the sequence $u^{(\alpha)}$ satisfies uniform bounds*

$$\max_P u^{(\alpha)} \leq C_0.$$

$$|V^{(\alpha)}|_{\mathrm{Euc}} \leq C_1.$$

Fix any $\delta > 0$. There are D_δ, s_δ with the following property. If p is a point in $P^{(\alpha)}$ with $s(p) \leq s_\delta$ and the distance of $p + s(p)\nu$ from the vertices of $P^{(\alpha)}$ is greater than δ then $D(p) \leq D_\delta$.

Here $V^{(\alpha)}$ is the vector field associated to $u^{(\alpha)}$ by the formula (5). Strictly speaking we should write $s_\alpha(p)$ etc., since these quantities depend on $P^{(\alpha)}$, but we hope that the meaning is clear. The arguments in this section do not depend strongly on the L^∞ bound on u, as opposed to a L^1 bound on the boundary value. The former is only used once, in the proof of Proposition 3, and could be avoided with a little extra work. Of course, we showed in the previous Section that the two conditions are in fact equivalent.

We will now explain the main idea of the proof of Theorem 3. We suppose, on the contrary, that there is a sequence of points p_α and $D(p_\alpha) = D_\alpha$ tends to infinity. Let us also suppose that p_α is the "worst" point, maximising the function D for each fixed α and that the sequence p_α stays a definite distance from the vertices. Then by performing a sequence of affine transformations, adding suitable affine linear functions and multiplying by D_α^{-1} we can obtain a sequence of convex functions $u_\flat^{(\alpha)}$ defined on large convex subsets of a half-plane $\{x_1 \geq 0\}$ with $u_\flat^{(\alpha)}(0,0) = 1$ and $u_\flat^{(\alpha)}$ attaining its minimum 0 at the point $(1,0)$. Our overall strategy is to obtain a contradiction by showing that the $u_\flat^{(\alpha)}$ have a C^0 limit and making various arguments with this, using the fact that p_α is the "worst" point. Two of the issues we have to deal with are

- In reality we need to use a more complicated definition of "worst" point, because of the constraint involving δ.
- We have to contend with the affine invariance of the problem, in choosing the affine transformations to define $u_\flat^{(\alpha)}$ appropriately. This is the choice of the parameter λ below.

The vector field V associated to the convex function u comes in to our arguments at a number of places and we will now recall two relevant points of theory. The first is that the vector field encodes the boundary conditions. Expressed in terms of coordinates, this says that the normal component of V at a point of an edge is fixed by the given measure σ. The second is that if we write $L = \log\det u_{ij}$ and introduce Legendre transform coordinates $\xi_i = \frac{\partial u}{\partial x_i}$ then

$$V^i = \frac{\partial L}{\partial \xi_i}. \tag{19}$$

This leads to a basic principle which will be important in our arguments. Suppose we have a bound on one component of the vector field: $|V^2| \leq C$ say. Then if, over a portion Γ of a contour $\{\xi_1 = \text{constant}\}$ the partial derivative $\frac{\partial u}{\partial x_2}$ varies by a bounded amount b say, then the ratio

$$\frac{\det u_{ij}(\gamma)}{\det u_{ij}(\gamma')}$$

is bounded by e^{Cb} for any $\gamma, \gamma' \in \Gamma$.

It is useful to have in mind a standard model for the boundary behaviour given by the function

$$u_0(x_1, x_2) = x_1 \log x_1 + x_2^2,$$

on the half-plane $\{x_1 > 0\}$. The associated vector field V has components $V^1 = 1, V^2 = 0$, and the function satisfies (1) with $A = 0$. Then in this case $D(p)$ is equal to 1 for all p. Now apply an affine transformation and, for $a \in \mathbf{R}$, set

$$u_a(x_1, x_2) = x_1 \log x_1 + (x_2 - ax_1)^2.$$

This also satisfies (1) with $A = 0$, and the same boundary conditions. At the point $p = (1, 0)$ we have $D(p) = a^2 + 1$. Since a can be made arbitrarily large, this shows that there is no way to derive an *a priori* estimate for $D(p)$ using only "local" information. However in this example we have $V^1 = 1, V^2 = a$, so the parameter a is detected by the tangential component of the vector field. This may help in understanding our main proof below, which uses the bound on V—obtained in our application by global arguments–to help control the quantity $D(p)$.

With this outline of the strategy in place, we now proceed in more detail. We suppose the polygon P has an edge given by

$$E = \{(x_1, x_2) : x_1 = 0, a \leq x_2 \leq b\},$$

and that P lies in the half-plane $\{x_1 > 0\}$. We can suppose that the measure $d\sigma$ on this edge is the standard Lebesgue measure dx_2. We consider a solution u of our equation (1) , where $|A| \leq C_2$. We fix a $\delta > 0$ and consider points $p = (s, t)$ with $s \leq s_\delta$ and $a + \delta \leq t \leq b - \delta$. We assume a bound on the boundary integral, as in the statement of Theorem 3. It is then easy to see that we can choose s_δ such that all such points p lie in P and that $sD(p)$ satisfies a fixed bound

$$sD(p) \leq C_3. \tag{20}$$

Similarly we have an elementary bound on the derivatives in the "tangential direction": for any two points p, p' satisfying the conditions above

$$|\frac{\partial u}{\partial x_2}(p) - \frac{\partial u}{\partial x_2}(p') \leq C_4. \tag{21}$$

Given s, t as above, we define the convex function u^* by normalising u at $p = (s, t)$. That is, u^* is given by adding an affine-linear function to u and u^* attains its minimum value 0 at p. Then, by definition,

$$D(p) = s^{-1} u^*(0, t).$$

Recall that the Guillemin boundary conditions around a vertex imply that the integral of the second derivative $u_{22} = \frac{\partial^2 u}{\partial^2_{x_2}}$, evaluated along the edge, *diverges*

at each end point. Let $\lambda_0 = \min(t-a, b-t)$ and for $\lambda < \lambda_0$ define

$$I(\lambda) = \int_{t-\lambda}^{t+\lambda} u_{22}(\tau, 0) d\tau.$$

Then $I(\lambda)$ is an increasing function, equal to 0 when $\lambda = 0$ and tending to infinity as $\lambda \to \lambda_0$. The same is true of $\lambda I(\lambda)$ so there is a unique λ such that

$$\lambda I(\lambda) = sD(p)/2.$$

The reason for this choice will appear presently.

The next proposition will be used to handle the potential "end-point" difficulties alluded to above.

Proposition 3 *There is a constant c, depending only on $\delta, C_0, C_2, C_3, C_4, \|A\|_{L^\infty}$ and the geometry of the polygon P, such that $s(p) \leq c\lambda^{-2}$.*

We need an elementary lemma.

Lemma 5 *There exists $\kappa > 0$ with the following property. Let f be any positive convex function on $[-1, 1]$ with $f(0) = 1$ and with $\int_{-1}^{1} f''(t) dt = 1$. Let $\mathcal{C}_f$ be the set of affine-linear functions σ such that*

- *$\sigma(t) \leq f(t)$ for all $t \in [-1, 1]$*
- *either $\sigma \leq 0$ on $[-1/2, \infty)$ or $\sigma \leq 0$ on $(-\infty, 1/2]$.*

Define $g(t) = \sup_{\sigma \in \mathcal{C}_f} \lambda(t)$. Then

$$\int_{-1}^{1} f - g \ dt \geq \kappa.$$

We leave the proof as an exercise.

To prove the Proposition we claim first that if $D(p)$ is large then s/λ is small. For if $\lambda \leq \delta/2$, then

$$I(\lambda) \leq \int_{t-\lambda}^{t+\lambda} u_{22}(0, \tau) d\tau \leq C_4.$$

So $sD(p) = \lambda I(\lambda) \leq \lambda C_4$ and $s/\lambda \leq C_4/S$.

On the other hand if $\lambda > \delta/2$ then

$$s/\lambda < 2s/\delta \leq 2C_3/D(p).$$

Now let W be the wedge-shaped region

$$W = \{(x_1, x_2) : |x_2 - t| \leq \frac{\lambda}{2s}(x_1 - s),$$

and let X be the intersection of P with W. It is clear that, when s/λ is small the centre of mass of P lies in X. Since X is convex, this implies that $\underline{u}_X \geq 0$, where $\underline{u}_X$ is the function defined in Section 2. So $|u - u_X| \leq u$ and

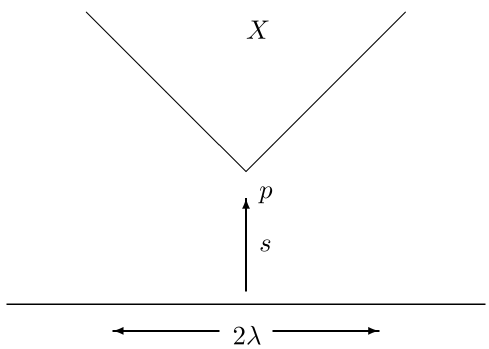
$$\int_{P \setminus X} (u - u_X) A \leq \|A\|_{L^\infty} \max u \ \ \text{Area } (P \setminus X). \tag{22}$$

Using our bound on $\max u$ we see that this integral is bounded by a multiple of the area of $P \setminus X$. So (8) gives

$$\int_{\partial P} u - u_X \leq c_1 \text{Area } (P \setminus X),$$

for some fixed c_1. It is also clear that the area of $P \setminus X$ is bounded by a multiple of s/λ, with the multiple depending only on the geometry of P. So we have

$$\int_{\partial P} u - \underline{u}_X \leq c_2 \left(\frac{s}{\lambda} \right). \tag{23}$$



We now make a similar argument to that in Lemma 2. Let q be a point of X, not equal to the centre of mass of P, and let π be the affine-linear function defining the supporting hyperplane to u at q. Thus the zero set of π is a line in the plane. Since $\pi(q) > 0$ and $\pi(p) < 0$ this line separates the points q and p. It follows that the restriction of π to the x_2 axis is negative on either the interval $(-\infty, -\lambda/2]$ or on the interval $[\lambda/2, \infty)$. Lemma 5 (after suitable rescaling) implies that

$$\int_{t-\lambda}^{t+\lambda} (u - \underline{u}_X)(0, \tau) d\tau \geq \kappa D(p) s \lambda.$$

So from (23) we obtain

$$\kappa D(p) \lambda \leq c_2 \frac{s}{\lambda},$$

which gives

$$D(p) \leq \frac{c_2}{\kappa} \lambda^{-2},$$

as required.

For t in the interval $(a+\delta, b-\delta)$ define

$$\Lambda(t) = \min(|(a+\delta) - t|, |(b-\delta) - t|).$$

Set

$$\mu = \max_{s,t} \Lambda(t)^2 D(s,t),$$

where t runs from $a+\delta$ to $b-\delta$ and s runs from 0 to s_δ. Choose a point (s_0, t_0) where the maximum is attained. Of course, this is the concept of the "worst point" alluded to above. Write $\lambda_0 = \lambda(s_0, t_0)$ and $D_0 = D(s_0, t_0)$. By the preceding Proposition we have

$$\lambda_0 \leq c\sqrt{\mu}\Lambda(t_0). \tag{24}$$

In other words if, as we suppose, μ is large the "scale" λ_0 in the x_2 direction is small compared with the distance to the end points $a+\delta, b-\delta$. Also for any fixed R we have

$$D(s', t') \leq D_0 \frac{1}{1 - R\sqrt{c}/\sqrt{\mu}} \tag{25}$$

for any t' with $|t_0 - t'| \leq R\lambda_0$.

We now define a new convex function $u_\flat$ by rescaling. We set

$$u_\flat(x_1, x_2) = \frac{1}{D_0 s_0} u^*(s_0 x_1, t_0 + \lambda_0 x_2). \tag{26}$$

Thus $u_\flat$ is defined on a polygon $P_\flat$ in the half-space $\{x_1 > 0\}$, which depends on P, s_0, t_0, λ_0. This polygon contains a rectangle

$$Q = \{0 < x_1 < cs_0^{-1}, |x_2| < c\lambda_0^{-1}\}$$

for some fixed c depending on P and δ. Since we are supposing that D_0 is large, both λ_0 and s_0 are small (by (20) and Proposition 3). Thus the rectangle is very large. By construction $u_\flat$ attains its minimum value 0 at the point $(1,0)$ and $u_\flat(0,0) = 1$. The choice of λ transforms into the condition that

$$\int_{-1}^{1} (\ u_\flat\ )_{22}(0,t)dt = 1/2. \tag{27}$$

The function $u_\flat$ satisfies an equation

$$\sum \frac{\partial^2 u_\flat^{ij}}{\partial x_i \partial x_j} = -A_\flat,$$

in $P_\flat$, where $A_\flat(x_1, x_2) = D_0 s_0 A(s_0 x_1, t_0 + \lambda_0 x_2)$. Thus, by (20),

$$|A_\flat| \leq C_3 \|A\|_{L^\infty}. \tag{28}$$

The function $u_\flat$ satisfies Guillemin boundary conditions along the edge $\{x_1 = 0\}$ but with the measure $D_0 dx_2$. This means that the normal component $V_\flat^1$ of the vector field associated to $u_\flat$ is $-D_0$ on the edge. Our overall goal is, roughly speaking, to show that this is impossible if D_0 is very large.

We finish this subsection with an observation which will be crucial in our proofs below. For any two points in Q the ratio of the determinant $\det(u_\flat)_{ij}$ evaluated at these two points is the same as the same as the ratio of $\det u_{ij}$ at the corresponding points in P. By (21) and the observation above we see that

for any two points q, q' on the intersection of any contour $\{\frac{\partial u_\flat}{\partial x_1} = \text{constant}\}$ with the large rectangle Q we have

$$\frac{\det(u_\flat)_{ij}(q)}{\det(u_\flat)_{ij}(q')} < \exp(C_1 C_4). \tag{29}$$

In the same vein, taking account of the rescaling, we get

$$(V_\flat)_2(x_1, x_2) = \frac{D_0 s_0}{\lambda_0} V^2(s_0 x_1, t_0 + \lambda_0 x_2),$$

so

$$|V_\flat^2| \leq \frac{D_0 s_0}{\lambda_0} C_1.$$

By the definition of λ_0 and (27) we have

$$\frac{D_0 s_0}{\lambda_0} = \int_{t_0 - \lambda_0}^{t_0 + \lambda_0} u_{22}(\tau, 0) d\tau \leq C_4,$$

so we have a uniform bound on $V_\flat^2$.

3.2 Boundedness

Applying the procedure above we obtain a sequence of convex functions $u_\flat^{(\alpha)}$ defined on large polygons in the half-space. In this subsection we show that these are bounded on compact subsets. As before, we usually omit the index α from the notation. Also, since we are only concerned with compact subsets we can be rather vague about the precise domains of definition of the functions.

The first point is that, by (25), and the scalings chosen we have, for any fixed compact subset, a bound

$$D(u_\flat; p) \leq k \tag{30}$$

for points in the set, where we can suppose k is as close to 1 as we please.

Consider the restriction of $u_\flat$ to the x_1-axis, and set $f(x) = u_\flat(x, 0)$. The maximising condition translates into the condition that

$$x f'(x) - f(x) + 1 \leq kx.$$

By construction, $f(1) = f'(1) = 0$ and we can integrate the differential inequality to obtain

$$f(x) \leq kx \log x.$$

This discussion is valid for x less than s_δ / s_0, which we know is large. The normalisation conditions $u_\flat(0,0) = 1$ and (27) easily imply that $u_\flat(0, x_2) \leq 10$ (say), for $-1 \leq x_2 \leq 1$. Now the convexity of $u_\flat$ yields a upper bound

$$u_\flat(x_1, x_2) \leq U(x_1, x_2),$$

say, for a suitable fixed function U and all points (x_1, x_2) in the triangle with vertices $(0,1), (0,-1), (s_\delta/s_0, 0)$. When we take our sequence $u_\flat^{(\alpha)}$ the apex s_δ/s_0 of this triangle tends to infinity.

It is less easy to obtain bounds on $u_\flat$ outside the triangle above, and for this we use the bound on the vector field $V_\flat$. As above, it suffices to bound the function $u_\flat(0, x_2)$. We prove

Proposition 4 *There is an $\eta > 0$ such that if for some $\sigma > 0$ we have a uniform bound $u_\flat^{(\alpha)}(0, \pm\sigma) \leq U_\sigma$ then there is a uniform bound $u_\flat^{\alpha)}(0, x_2) \leq U_{\sigma+\eta}$ for all $|x_2| \leq \sigma + \eta$.*

Since η is fixed we can use this to bound the function $u_\flat$ on any compact subset of the half-plane.

Write ξ_i for the partial derivatives $\frac{\partial u_\flat}{\partial x_i}$. The first step in the proof of Proposition 4 is an elementary Lemma. We consider a disc Δ of small radius r centred on the point $(\frac{1}{4}, 0)$.

Lemma 6 *We can fix small r and a $Z > 0$ such that on the disc Δ we have*

$$\xi_1 < -\frac{1}{2} \quad , \quad -Z < \xi_2 < Z,$$

for large enough α.

We can choose the disc to lie well inside the triangle on which we have bounds on $u_\flat$, and this gives bounds on the derivatives ξ_i, by convexity. Thus the only

thing is to arrange that $\xi_1 < -\frac{1}{2}$ on Δ. For this we begin by choosing r so small that such that if $|x_2| < r$ we have

$$u_\flat(0, x_2) \geq .99 \ \ , \ u_\flat(1, x_2) \leq .01.$$

This is clearly possible. Consider a point (a_1, a_2) with $|a_2| < r$ and $a_1 < 1$ at which $\xi_1 = -1/2$. We show that a_1 cannot be too small. By (25) we have

$$u_\flat(0, a_2) \leq u_\flat(a_1, a_2) + a_1/2 + ka_1,$$

where k can be made as close to 1 as we like. So

$$u_\flat(a_1, a_2) \geq .99 - (k + .5)a_1.$$

On the other hand, convexity implies that

$$u_\flat(1, a_2) \geq u_\flat(a_1, a_2) - (1 - a_1)/2,$$

so

$$u_\flat(a_1, a_2) \leq .5 + .01 - a_1/2.$$

We deduce from these inequalities that $a_1 \geq .48k^{-1}$ so, taking k close to 1 and r small, such a point cannot lie in Δ. This establishes the Lemma.

We now fix the disc Δ and the number Z, as above. Let τ_t be the flow by translations on $\mathbf{R}^2$

$$\tau_t(\xi_1, \xi_2) = (\xi_1, \xi_2 + t),$$

and let Ψ_t be the corresponding flow on the polygon $P_\flat$. That is

$$\Psi_t = (Du_\flat)^{-1} \circ \tau_t \circ (Du_\flat).$$

We consider the images $\Psi_t(\Delta)$, for parameters $t \in [0, T]$. Our general principle tells us that, if these are all contained in the large rectangle Q, then the area of each such image is at least a fixed multiple of the area of Δ. (Using the fact that the Jacobian of $Du_\flat$ is $\det(u_\flat)_{ij}$.)

Now suppose we have a $\sigma > r$ for which we have obtained an *a priori bound* $u_\flat(0, a_2) \leq U$, for all $|a_2| \leq \sigma$. Consider a point (x_1, x_2) where $\xi_1 \leq -1/2$. We first treat the case when $|x_2| \leq \sigma$. Then we have, by convexity and the positivity of $u_\flat$,

$$U \geq u_\flat(0, x_2) \geq u_\flat(x_1, x_2) + x_1/2 \geq x_1/2. \tag{31}$$

Second we treat the case when $x_2 = \sigma + t$ where $0 \leq t \leq \eta$, and η will be chosen shortly. By convexity we have

$$u_\flat(x_1, \sigma + t) - (x_1\xi_1 + t\xi_2) \leq u_\flat(0, \sigma) \leq U. \tag{32}$$

By (25) we have

$$u_\flat(0, \sigma + t) \leq u(x_1, \sigma + t) - \xi_1 x_1 + kx_1.$$

The two together give

$$u_\flat(0,\sigma+t) \leq U + t\xi_2 + kx_1. \tag{33}$$

On the other hand, the first inequality gives

$$u_\flat(x_1,\sigma+t) \leq U + (x_1\xi_1 + t\xi_2),$$

and since $u_\flat(x_1,\sigma+t) \geq 0$ we have

$$x_1\xi_1 \geq -(U + t\xi_2). \tag{34}$$

Since $\xi_1 \leq -1/2$ we obtain

$$x_1 \leq 2(U + t\xi_2). \tag{35}$$

Substitute back into (33) to get

$$u_\flat(0,\sigma+t) \leq (1+2k)(U + t\xi_2). \tag{36}$$

Now let ρ be the minimum value of ξ_2 on the set where $x_2 = \sigma + \eta$ and $\xi_1 \leq -1/2$. By the above we have

$$u_\flat(0,\sigma+\eta) \leq (1+2k)(U + \eta\rho),$$

so we want to show that ρ is not too large. Let S be the union of the rectangle $\{0 \leq x_1 \leq 2U, -r \leq x_2 \leq \sigma\}$ and the rectangle $\{0 \leq x_1 \leq (1+2k)(U+\eta\rho), \sigma \leq x_2 \leq \sigma+\eta\}$. We assume the hypothesis that S lies in the large rectangle Q. By the inequalities (30),(35) above; for any t with $0 \leq t \leq \rho - Z$ the image $\Psi_t(\Delta)$ under the flow lies in S. If $|t - t'| \geq 2Z$ the images $\Psi_t(\Delta), \Psi_{t'}(\Delta)$ are disjoint. Since the area of each of these is at least a fixed multiple of the area of Δ we deduce that

$$\rho \leq \kappa \text{Area } (S),$$

for some κ which does not depend on U. (This constant κ depends only on r, Z, C_1, C_4.) Writing down the area of S, we get

$$\rho \leq \kappa \left(2(r+\sigma)U + \eta(1+2k)(U+\eta\rho)\right). \tag{37}$$

Choose η so that $(1+2k)\kappa\eta^2 < 1/2$, say. Then we can rearrange (37) to get an upper bound on ρ, which then gives by (36) a bound on $U(\sigma+\eta)$. (Of course, we use a symmetrical argument for negative values of x_2.) The final detail to add in the argument concerns our hypothesis that S lies in the large rectangle. But, assuming this hypothesis, the set actually lies in a smaller rectangle (because of the bound on ρ) so we can easily establish the truth of the hypothesis, for large α, by a continuity argument.

3.3 The blow-up limit

Let us review the argument of this section thus far. We suppose we have a sequence $u^{(\alpha)}$ of solutions corresponding to data sets $P^{(\alpha)}, A^{\alpha)}, \sigma^{(\alpha)}$, satisfying the conditions of Theorem 3 and with μ_α tending to infinity. We want to obtain a contradiction. We rescale to get convex functions $u_\flat^{(\alpha)}$ defined on a sequence of domains $P_\flat^{(\alpha)}$ which exhaust the half-plane $\{x_1 > 0\}$ and we have shown that these are bounded on compact subsets of the closed half-plane. After taking a subsequence, we can suppose that the $u_\flat^{(\alpha)}$ converge uniformly on compact subsets of the *open* half-plane to a limit $u_\flat^{(\infty)}$, which is a continuous, weakly convex, function on the half-plane. The main idea of our proof is to obtain a contradiction by an analysis of this limit. First, in the next Proposition, we show that it cannot be strictly convex anywhere. The proof we give uses the assumed bound on the associated vector fields. The author knows of other arguments, for this step, which avoid that assumption, but which are longer.

Recall that a convex function v is called strictly convex at a point p if there is an affine linear support function π such that $v - \pi$ has a unique minimum at p.

Proposition 5 *The limit $u_\flat^{(\infty)}$ is not strictly convex at any point of the half-plane.*

Suppose the contrary, so there is a small disc D in the half-plane centred at a point p and an affine-linear function π such that $u_\flat^{(\infty)} - \pi$ vanishes at p but is strictly positive on the boundary of D. The smooth functions $u_\flat^{(\alpha)}$ satisfy elliptic equations

$$\left(u_\flat^{(\alpha)}\right)^{ij}_{ij} = -A^{(\alpha)}$$

with a fixed bound on $\|A^{(\alpha)}\|_{L^\infty}$. If we set Also $u_\flat^{(\alpha)}(p) - \lambda(p) \to 0$ and $u_\flat^{(\alpha)} - \lambda \geq \delta > 0$ say on ∂D, once α is sufficiently large. Now these facts give complete control of the functions $u_\flat^{(\alpha)}$ in the interior of D. We can apply Theorem 5 from [5] to obtain upper and lower bounds on the Jacobians $\det((u_\flat^{(\alpha)})_{ij})$ over the interior of D and arguing as in [4], using the theory of Cafarelli and Gutierrez, bootstrap to control all higher derivatives. In particular the vector fields $V_\flat^{(\alpha})$ defined by

$$((V^{(\alpha)})^i = -\left(u_\flat^{(\alpha)}\right)^{ij}_j,$$

are uniformly bounded on a small neighbourhood of p.

To obtain a contradiction, suppose $p = (p_1, p_2)$ and consider a rectangle

$$S = \{(x_1, x_2) : |x_2 - p_2| \leq \eta, 0 < x_1 < p_1\},$$

with η small. Since the divergence of $V_\flat^{(\alpha)}$ is bounded the total flux of $V_\flat^{(\alpha)}$ through the boundary of S is small. The flux through the two edges where

$x_2 = p_2 \pm \eta$ is bounded, by our bound on the x_2 component of $v^{(\alpha)}$ and the flux through the edge where $x_1 = p_1$ is bounded by the argument of the previous paragraph. But the boundary conditions, after rescaling, imply that the x_1 component of $V_\flat(\alpha)$ along the remaining edge, in the boundary of the half-plane, is D_α, so the flux through this edge is $2\eta D_\alpha$ which tends to infinity by hypothesis.

For $a \in \mathbf{R}$ let Γ_a denote the set of points (x_1, x_2) where $x_1 > 0$ and $x_1 = 1 + ax_2$. This is either a half-line or, in the case when $a = 0$, a line.

Corollary 3 *The limit $u_\flat^{(\infty)}$ vanishes on a set Γ_a for some $a \neq 0$.*

To see this let Z be the zero set of $u_\flat^{(\infty)}$ in the open upper half-plane. Recall that $u_\flat^{(\alpha)}$ is normalised to achieve its minimum at the point $p_0 = (1, 0)$. Thus Z is a convex set containing p_0. Proposition 5 implies that Z has no extreme points and it follows immediately that there must be a line through p_0 whose intersection with the upper-half plane is contained in Z. We know that $u_\flat^{(\alpha)}(t, 0)$ is bounded below by $t \log t - t + 1$ for $0 < t < 1$ and it follows that Z cannot contain the line segment $\{x_2 = 0, x_1 > 0\}$. Thus Z contains Γ_a for some a and it only remains to rule out the possibility that $a = 0$. To do this, recall that we chose our normalisation so that the x_2 derivative of $u_\flat^{(\alpha)}$ differs by 1 at the two points $(0, \pm 1)$. This obviously implies that we can find some fixed c such that $u_\flat^{(\alpha)}(0, c) > 2$ say. Suppose that Z contains Γ_0, so $u_\flat^{(\alpha)}(1, c)$ tends to zero as α tends to infinity. It follows that there must be some sequence b_α tending to 1 such that the x_1 derivatives of $u_\flat^{(\alpha)}$ evaluated at (c, b_α) converge to 0. But then these points contradict (30), once $k < 2$.

3.4 The final contradiction

We again pause for discussion. Changing our coordinates slightly, we may without real loss of generality suppose that $a = 1$ and $u_\flat^{(\infty)}$ vanishes on the ray $\{x_1 = x_2 + 1, x_1 > 0\}$. The essential case to have in mind is when $u_\flat^{(\infty)} = \max(x_2 - x_1 + 1, 0)$ so let us momentarily assume that we have this case. It is tempting to try to argue as follows. For any large C and large enough α we can find a point p' near to $(C + 1, C)$ such that $D(p', u_\flat^{(\alpha)}$ is very close to 1. In other words, transferring back to the original functions $u^{(\alpha)}$ there are points much further from the edge than the "worst point" but which are "almost as bad". So this strongly suggests that if D becomes large close to the boundary of P it must also become large in the interior, which is ruled out by our hypotheses. Indeed if we were to drop the hypothesis on the integral over the boundary then we would see exactly this phenomenon, as we discuss further in Section 6. However, while it is suggestive, it seems hard to turn this line of argument into an actual proof. The proof we give below is rather different and hinges on our general principle that $J = \det(u_\flat^\alpha)_{ij}$ changes by a bounded factor

on the parts of the contours $\{\xi_1 = \text{constant}\}$ in the large rectangle Q. We will show that this leads to a contradiction. The argument is similar to that used to prove Proposition 4 above.

Lemma 7 *The boundary values $u_\flat^{(\alpha)}(0,x_2)$ converge to $x_2 - 1$ uniformly for x_2 in any closed interval $[-1,R]$.*

For given $x_2 > -1$ we can find a sequence $x_1^{(\alpha)}$ converging to $1 + x_2$ such that the values $u_\flat^{(\alpha)}$ and the partial derivatives $\frac{\partial u_\flat^{(\alpha)}}{\partial x_1}$ evaluated at $(x_1^{(\alpha)}, x_2)$ converge to 0 as $\alpha \to \infty$. Then we obtain from (30) that

$$u_\flat(0,x_2) \leq k_\alpha(x_2+1) + \epsilon_\alpha$$

where $\epsilon_\alpha \to 0, k_\alpha \to 1$. By construction, $u_\flat^{(\alpha)}(0,0) = 1$ and it follows from convexity that $u_\flat(0,x_2)$ tends to x_2+1, uniformly for x_2 in any compact subset of $(-1,\infty)$. However the functions are bounded on a neighbourhood of the point $(0,-1)$ and it follows again from convexity that the convergence is uniform up to $x_2 = -1$.

Given a small number r consider the region

$$\Omega = \{(x_1,x_2) : x_1 > 0, r < \sqrt{|x_1|^2 + |x_2+1|^2} < r^{-1}\}.$$

We consider first the points in Ω where the partial derivative ξ_1 of $u_\flat^{(\alpha)}$ is $-1/2$ (say) and ξ_2 is $1/10$ (say). To simplify the exposition imagine first that $u_\flat = u_\flat^\alpha$ *vanishes* on the intersection of Ω with the ray. Thus ξ_1, ξ_2 also vanish on this set. Suppose $\xi_1 = -1/2$ and $x_2 > -1$. Then we must have $x_1 < x_2 + 1$ and

$$0 = u_\flat(1+x_2,x_2) \geq u_\flat(x_1,x_2) - \frac{1}{2}(1+x_2-x_1).$$

On the other hand

$$(x_2+1) - \epsilon_\alpha \leq u_\flat(0,x_2) \leq u_\flat(x_1,x_2) + (k_\alpha + \frac{1}{2})x_1,$$

where $\epsilon_\alpha \to 0$. These imply that

$$\frac{1}{2}(x_2+1) - kx_1 \leq \epsilon_\alpha.$$

On the other hand, just from the fact that $u_\flat(x_1,x_2) \geq 0$ we have

$$0 \leq u_\flat(0,-1) + (x_1\xi_1 + (x_2+1)\xi_2).$$

So if $\xi_1 = -1/2, \xi_2 = 1/10$ we have

$$x_1 \leq 2u_\flat(0,-1) + (x_2+1)/5.$$

If $u_\flat(0,-1)$ and ϵ_α are sufficiently small then these inequalities have no common solution in Ω. It is clear from a continuity argument then that the point where

$\xi_1 = -1/2, \xi_2 = 1/10$ must lie in the small half-disc D of radius r about the origin.

Now obviously the same argument applies to values of ξ_1, ξ_2 close to $-1/2, 1/10$. Further, it is easy to extend the argument to the case when $u_\flat$ is C^0 close to a function vanishing along the ray, over the fixed annulus. So we conclude that there is a small rectangle $R \subset \mathbf{R}^2$ of the form

$$R = \{(\zeta_1, \zeta_2) : |\zeta_1 + 1/2| < \eta, |\zeta_2 - 1/10| < \eta\}$$

with the following property. For any given r and all large enough α all points (x_1, x_2) for which which $(\xi_1(x_1, x_2), \xi_2(x_1, x_2)) \in R$ are contained in D.

Now for fixed large α take a point (ζ_1, ζ_2) in R and consider the contour $\xi_1(x_1, x_2) = \zeta_1$. This contour meets the line $x_2 = 0$ at some point $(a_1, 0)$. As in the proof of Proposition 4 we have $a_1 > c$ for some fixed $c > 0$. Then using the estimate in Theorem 5 of [5] we have an upper bound on the determinant function at this point. It is obvious that when $x_1 > -1$ the contour cannot move out of the large rectangle Q. We conclude from our principle that the determinant is bounded at all points whose derivative lies in R, say $J \leq C$. But the inverse of the derivative maps R into D so

$$\int_R J^{-1} d\zeta_1 d\zeta_2 \leq \text{Area } (D) = \pi r^2/2.$$

Thus $4C^{-1}\eta^2 \leq \pi r^2$. But since r can be made arbitrarily small, with η fixed, this gives our contradiction.

4 C^∞ limits away from the vertices

In the previous section we obtained a uniform bound on the quantity $D(p; u)$ along the interior of each edge. We now use this to get complete control of the solution away from the vertices. This is straightforward, given the results from [5], if we have a lower bound on the Riemannian distance function determined by the solutions, and we explain this argument in (4.1). The main work of the section goes in to establishing this lower bound. For this we derive various estimates on the solution, near to an edge, and particularly on $\det u_{ij}$. These estimates may have independent interest.

4.1 The proof, assuming a lower bound on the Riemannian distance

We begin with the relation between the quantity $D(p; u)$ and the "M-condition". Recall that in [5] we said that u satisfies an M condition if $V(p, q) \leq M$ for any pair p, q of points in P such that the line segment $\{tp + (1-t)q : -1 \leq t \leq 2\}$ lies in P. Here $V(p, q)$ is the variation of the derivative of u in the direction of the unit vector $\nu = (p-q)/|p-q|$ between the two points. For brevity we will call such pairs p, q "admissible pairs".

Proposition 6 *Suppose u is a normalised function on the polygon P and we have*

- *A bound on the integral of u over ∂P;*
- *For each $\delta > 0$ a bound on $D(p; u)$ for points p whose Euclidean distance to all vertices of P exceeds δ.*

Then for any $\delta' > 0$ the variation $V(p, q)$ is bounded for all admissible pairs p, q where the Euclidean distance from p to the vertices exceeds δ'.

This is very elementary, so we will use rather informal language. The first hypothesis controls $V(p, q)$ when p is not close to the boundary, so the relevant case is when p is close to a unique edge. We suppose, as in the previous section, that this edge is a segment $a \leq x_2 \leq b$ of the x_2-axis and $p = (p_1, p_2)$ with $a + \delta < p_2 < b - \delta$. Let u_p be the function obtained from u by normalising at p. As in Proposition 4, the bound on $D(p')$, for points p' on the segment $\{p'_2 = p_2\}$, gives a bound

$$u_p(p'_1, p_2) \leq C p_1, \tag{38}$$

for $0 \leq p'_1 \leq 3p_1$, say. With this point p, the points q we need to consider in the definition of the M-condition range over some quadrilateral Q. Two of whose edges are segments in the lines $\{x_1 = p_1/2\}, \{x_1 = 2p_1\}$ and the other two are determined by the other edges of P. But these other two edges are a definite distance from the rest of the boundary of P. We can choose a slightly larger quadrilateral Q^+, two of whose edges are segments in the lines $\{x_1 = 0\}, \{x_1 = 2p_1\}$ and whose other two edges are again a definite distance from the rest of the boundary of P. For each unit vector ν there are unique h, h^+ such that $q = p + h\nu$ lies in the boundary of Q and $q^+ = p + h^+\nu$ lies in the boundary of Q^+. We can suppose that $h \leq (1 - \epsilon)h^+$ for some fixed $\epsilon > 0$. Then if $q = p + h\nu$ we have

$$V(p, q) \leq \frac{u_p(q^+)}{\epsilon h^+}.$$

Write the x_2 coordinate of q^+ as $p_2 + t$. When $t = 0$ then (38) states that $u_p(q^+) \leq C p_1$. When q^+ lies on one of the other two edges of Q^+ (not parallel to the x_2-axis) we have a bound $u_p(q^+) \leq C$, since these edges are a definite distance from the other edges of P. Convexity of u_p yields an inequality of the form $u_p(q^+) \leq C(p_1 + |t|)$. On the other hand we have $h^+ \geq C\sqrt{p_1^2 + t^2}$ so

$$\frac{u_p(q^+)}{h^+} \leq C \frac{p_1 + |t|}{\sqrt{p_1^2 + t^2}},$$

which is bounded. This completes the proof.

Let $\Omega \subset \overline{P}$ be the set obtained by deleting fixed small Euclidean discs about the vertices and let $\partial_* \Omega$ be that part of the boundary of Ω which is

not contained in the boundary of P. An admissible convex function u on P defines a Riemannian metric u_{ij}on $\overline{P}$, regarded as a 2-manifold with corners. For $p, q \in \Omega$ we write $\text{dist}_u(p, q)$ for the Riemannian distance defined by this metric, and

$$\text{dist}_u(p, \partial_* \Omega) = \inf_{q \in \partial_* \Omega} \text{dist}_u(p, q).$$

Locally, we may also associate a 4-dimensional Riemannian manifold to this data and we let $|F|^2$ be the square of the Riemannian norm of the curvature tensor as in [4], [5]. Now consider a sequence $u^{(\alpha)}$ as in Theorem 3. We claim that

Proposition 7 *There is a fixed bound* $|F^\alpha|^2 \mathbf{dist}(\ ,\partial\Omega)^2 \leq C$ *for all* α.

Of course here, strictly speaking we have a sequence of polygons $P^{(\alpha)}$ so we need to fix a sequence of domains $\Omega^{(\alpha)}$, but the meaning should be clear.

The proof of Proposition 7 is a straightforward modification of the arguments of [5], which we only outline. We proceed by contradiction and suppose there is a sequence of points p_α for which $K_\alpha = |F|^2 \mathbf{dist}(\ ,\partial_*\Omega)^2$ tends to infinity. Then we rescale the metric so that after rescaling the curvature has norm 1 at the chosen points. After this rescaling the distance to the boundary $\partial_*\Omega$ is K_α which becomes large by hypothesis, and the curvature is bounded on balls of a fixed size about the chosen points. This means that we can take the blow-up limit just as in [5] and the extra boundary "disappears" in the limit. Then the analysis of the blow-up limits in [5] gives the desired contradiction. We have to use the M-condition a number of times in these arguments, but only at points in Ω, and we have this by Proposition 6.

Now fix a subset $\Omega_0 \subset \Omega$, for example removing larger Euclidean discs about the vertices. We will show

Proposition 8 *There is an* $\eta > 0$ *such that, for all* α

$$\text{dist}_{u^{(\alpha)}}(\partial_*\Omega_0, \partial_*\Omega) \geq \eta.$$

Assuming this, Proposition 7 gives an upper bound on the size of the curvature tensor over Ω_0 and the arguments of [5] apply without change to give C^∞ convergence of the $u^{(\alpha)}$, in the same sense as in [5]. Since we can make Ω, Ω_0 as large as we please, we conclude that the u_α converge in C^∞ on compact subsets of $\overline{P}$ minus the vertices. The proof of Proposition 8 takes up the remainder of this section.

4.2 Lower bound on Riemannian distance: strategy

By the results of [4] we know that over any compact subset of the open polygon P the Riemannian length of paths compares uniformly with the Euclidean length. Also we know that, given a bound on the quantities $D(p)$, the Riemannian length of a line segments meeting an edges in an interior points is bounded below ([5], Lemma 2). Using these facts, it is elementary to reduce the proof of Proposition 8 to the following. Given any two points q, q' in the interior of an

edge, there is a lower bound on the Riemannian length of paths from q to q'. We can take the edge to be a segment in the x_2-axis and $q = (0, \alpha), q' = (0, \beta)$, with $\alpha < \beta$. The same elementary arguments show that it suffices to consider paths which lie in a rectangle $\{(s,t) : 0 \leq s \leq s_0, \alpha \leq t \leq \beta\}$, for arbitrarily small s_0.

Remarks

1. Of course we fix s_0 so that this rectangle is well away from the other edges of P.

2. It is not hard to avoid appealing to the results of [4] here, at the cost of some extra arguments.

3. The obvious path, given by the line segment in the x_2-axis, between these points is a geodesic and we *expect* that this will be the length minimising path. If we knew this then the proof of Proposition 8 would be substantially simpler—we could avoid Proposition 10 below—but the author has not found a argument to establish this fact so we have to work more.

Our basic idea is to consider the function $\xi_2 = \frac{\partial u}{\partial x_2}$ on P. For a pair of points $(0, t_1), (0, t_2)$ on the edge, with $t_1 < t_2$, write

$$\Delta(t_1, t_2) = \xi_2(0, t_2) - \xi_2(0, t_1) = \int_{t_1}^{t_2} u_{22} \, dt.$$

One step in the proof is to establish that $\Delta(\alpha, \beta)$ is not small (Corollary 4 below). To see the relevance of this consider, for this exposition, the linear path along the x_2-axis. The Riemannian length of this path is

$$\int_{\alpha}^{\beta} \sqrt{u_{22}} \, dt,$$

while

$$\Delta(\alpha, \beta) = \int_{\alpha}^{\beta} u_{22} \, dt.$$

Informally, we expect that if $\Delta(\alpha, \beta)$ is not small then u_{22} should not be small at typical points and so the Riemannian length should not be small. But of course this argument does not suffice, as it stands, because of the square-root in the integral for the Riemannian length. Much the same issue arose in [4], deriving estimates in the interior of the polygon. The analogous difficulty there was to obtain lower bounds for the Riemannian distance to the boundary given a "strict convexity" condition. The approach in [4] was to replace Riemannian balls with "sections" of the convex function, using deep results of Caffarelli. The problem at hand is that we are working up to the boundary, where these results do not apply.

To proceed with our outline of the strategy, consider the square of the Riemannian norm of its derivative $\nabla\xi_2$ which is

$$|\nabla\xi_2|^2 = u^{ij}\sum \frac{\partial\xi_2}{\partial x_i}\frac{\partial\xi_2}{\partial x_j} = u^{ij}u_{2i}u_{2j} = u_{22}.$$

So along any path in P with the given end points the change in ξ_2 is bounded by $\int\sqrt{u_{22}}\, d\sigma$, where $d\sigma$ denotes Riemannian arc length along the path. Thus if we have an *upper* bound $u_{22} \leq C$ along the path, we have

$$\Delta(\alpha,\beta) \leq L\sqrt{C},$$

where L is the length of the path; so we have a lower bound $L \geq C^{-1/2}\Delta(\alpha,\beta)$, as desired. This upper bound on u_{22} is given in Proposition 10 below. The proof of this, and the lower bound on $\Delta(\alpha,\beta)$ goes through estimates for the determinant $J = \det u_{ij}$. We emphasise that in all of these arguments we make much use of the result of Section 4; $D(p;u) \leq D$ say, for all relevant points p, and the various constants in our statements depend on D.

4.3 Lower bound on Riemannian distance: detailed proofs

Lemma 8 *Let $t_1 < t_2$ be two points in the interval $[\alpha,\beta]$ and $\tau = (t_1+t_2)/2$. There are constants c,c' such that if for some $s \leq s_0$ we have $\Delta(t_1,t_2) \leq c\frac{s}{t_2-t_1}$ then $J(s,\tau) \leq c'(t_2-t_1)^{-2}$.*

The proof is sufficiently like Lemma 14 in [5] that we leave this to the reader. (Elementary arguments give bounds on the first derivative in a suitable neighbourhood of (s,τ), then we apply the maximum principle result Theorem 5 of [5].)

Corollary 4 *For any $\mu > 1$ there is a constant C_μ such that*

$$\Delta(t_1,t_2) \geq C_\mu(t_2-t_1)^\mu.$$

In particular $\Delta(\alpha,\beta)$ is bounded below.

To see this use Theorem 5 in [4] which states that for any $a < 1$ the function J satisfies a lower bound $J(s,t) \geq Cs^{-a}$. Then the statement follows immediately after re-arranging the inequalities. Note that if we could take $\mu = 1$ we would we in a strong position—u_{22} would then be bounded below on the edge— but the author has not been able to achieve this directly.

The next step is to find a sharp *upper* bound on the function J.

Proposition 9 *There is a constant C such that $J(s,t) \leq Cs^{-1}$ for all $t \in [\alpha,\beta]$.*

We choose nested intervals $(\alpha,\beta) \subset (\alpha',\beta') \subset (\alpha'',\beta'')$ so that the rectangle $(0,s_0] \times [\alpha'',\beta'']$ is well away from the other edges of P. We have an upper bound $\Delta(\alpha'',\beta'') \leq \Delta''$ say. We can suppose that the result of Lemma 8 applies

in the larger interval $[\alpha'', \beta'']$. We have an upper bound, $J(s_0, t) \leq J_0$ say, if $\alpha'' < t < \beta''$. Set $\eta_0 = J_0^{-1} s_0^{-1}$ and consider some η with $0 < \eta \leq \eta_0$.

Consider the function $F = J^{-1}$. This satisfies the equation $u^{ij} F_{ij} = -A < 0$ (see (14) in [4]). Thus the function $G = F - \eta x_1$ has no interior minima. We have $G = 0$ on the axis $\{x_1 = 0\}$ and $G > 0$ on the parallel segment $\{x_1 = s_0, \alpha'' < x_2 < \beta''\}$. Let Q be the rectangle $(0, s_0) \times (\alpha', \beta')$ and Σ be the subset of Q on which $G < 0$. Suppose Σ contains a point $p = (p_1, p_2)$ with $\alpha < p_2 < \beta$. Then the connected component of Σ containing p must meet the boundary of Q, since there are no interior minima and by construction this can only occur on the boundary components $x_2 = \alpha', \beta'$. So there is a continuous path in S from p to either the boundary $x_2 = \alpha'$ or to $x_2 = \beta'$. Without loss of generality suppose the former. Then for each $\tau \in (\alpha', \alpha)$ there is a point (s, τ) in Σ, *i.e.* where $J(s, \tau) > \eta^{-1} s^{-1}$. Now let $\lambda = \sqrt{c' \eta s}/2$, with c' as in Lemma 8. We suppose η is chosen so that $\sqrt{c' \eta s_0}/2 < \alpha' - \alpha''$ thus $\lambda < \alpha' - \alpha''$ and the interval $[\tau - \lambda, \tau + \lambda]$ is contained in $[\alpha'', \beta'']$. By Corollary 4 we have

$$\int_{\tau-\lambda}^{\tau+\lambda} u_{22} dt \geq c \frac{s}{\lambda} = \left(\frac{4c}{c' \eta} \right) \lambda. \tag{39}$$

Let f be the restriction of the second derivative u_{22} to the interval $[\alpha'', \beta'']$ in the edge, extended by zero to a function on $\mathbf{R}$. Thus $\| f \|_{L^1} \leq \Delta''$. Let m_f be the *maximal function* of f;

$$m_f(\sigma) = \max_{\mu > 0} \frac{1}{2\mu} \int_{\sigma-\mu}^{\sigma+\mu} f(t) \ dt.$$

Thus

$$\frac{1}{2\lambda} \int_{\tau-\lambda}^{\tau+\lambda} u_{22} dt \leq m_f(\tau).$$

and (39) gives $m_f(\tau) \geq \frac{2c}{2c'\eta}$ for each $\tau \in [\alpha', \alpha]$. Now the weak type bound for the maximal function tells us that there is a constant C such that for all b the measure of the set on which m_f exceeds b is at most $C \| f \|_{L^1} b^{-1}$. Thus

$$(\alpha - \alpha') \leq \frac{C \Delta'' c'}{2c} \eta.$$

If we choose η sufficiently small we get a contradiction, so there can be no such point p. In other words $J(s, t) \leq \eta^{-1} s^{-1}$ for $t \in [\alpha, \beta], s \leq s_0$.

It is easy to see, from the Guillemin boundary counditions, that the limit of $sJ(s, t)$ as $s \to 0$ is the second derivative u_{22}, evaluated at the point $(0, t)$. So a Corollary of the result above is that u_{22} is bounded on the interval $[\alpha, \beta]$ in the edge. This then gives us a lower bound on the Riemannian length of this interval and, as in the third remark at the beginning of Section 4.1, we strongly suspect that this actually realises the minimal length. However, lacking a proof of this, we go on to prove.

Proposition 10 *There is a constant C such that $u_{22} \leq C$ at all points (s,t) with $s \leq s_0, \alpha \leq t \leq \beta$.*

This result completes the proof of Proposition 8, as we explained in 4.1.

The proof of Proposition 10 is roughly speaking to argue that if u_{22} is large then J would violate the bound of Proposition 9.

Lemma 9 *Given $D > 0$ there are positive $\kappa, \zeta_1, \zeta_2, R > 1$ with the following property. Suppose v is a smooth convex function on th rectangle $\{0 \leq x_1 \leq R, -R \leq x_2 \leq R\}$, whose derivative is a diffeomorphism to its image. Write v_i for the partial derivatives $\frac{\partial v}{\partial x_i}$. Suppose that $v_1(x_1,x_2) \to -\infty$ as $x_1 \to 0$. Suppose that v satisfies a bound $D(p;v) \leq D$ for all points p. Suppose that v is normalised at the point $(1,0)$, that $v(1,x_2) \leq 2D$ for $-1 \leq x_2 \leq 1$ and $v(1,1) = 2D$. Then any point (x_1,x_2) where $v_1 \leq -\zeta_1$ and $\zeta_2 < v_2 < 2\zeta_2$ has $|x_1|, |x_2| \leq \kappa$.*

The proof of this is similar to the arguments in Proposition 4 and Lemma 7. All the steps are entirely elementary so we will use informal language. Given ζ_i, let S be the set of points with $v_1 \leq -\zeta_1$ and $\zeta_2 < v_2 < 2\zeta_2$. We choose $\zeta_2 \leq 1/2$ so the hypotheses imply that for any $\zeta \in [\zeta_2, 2\zeta_2]$ there is an x_2 in $(0,1)$ such that $v_2 = \zeta$ at the point $p_\zeta = (1, x_2)$. We consider the contour Γ on which $v_2 = \zeta$ and $x_1 \leq 1$. This cuts each line $\{x_1 = \text{constant}\}$ exactly once (so long as the intersection point does not move out to the boundary $x_2 = \pm R$). We have $v(x_1, 0) \leq D$ for $0 \leq x_1 \leq 1$. Then convexity implies that no point (x_1,x_2) with $0 \leq x_1 \leq 1$ and x_2 very negative can lie in Γ. Given a large positive ρ we consider the line through the points $(1,2)$ and $(0,\rho)$. We have an upper bound on the value of v at the intersection of this line with the x_1 axis while $v(1,2) = 2D$ by hypothesis. Convexity implies that at a point (x_1,x_2) on this line with $0 \leq x_1 \leq 1$ the value of v must be approximately $2Dx_2$. For suitable choices of the parameters we see that the contour Γ cannot meet this line segment. In particular, Γ is confined to lie in a bounded region $Q = \{0 \leq x_1 \leq 1, -\rho \leq x_<2 \leq \rho\}$ say. (We can suppose $R > \rho$ so we do not have any difficulties with the domain of definition.) As we move along the contour Γ, with x_1 decreasing, the derivative v_1 tends to $-\infty$ so whatever the value of ζ_1 the point on the contour eventually lies in S. On the other hand if we choose ζ_1 large then the bound on the $D(p)$ implies that p_ζ is *not* in S. The hypotheses imply that S is connected and it follows that S is contained in the bounded set Q, which completes the proof.

We now prove Proposition 10. Given a point $p = (s,t)$, we let u_p be the function obtained by normalising u at p. By applying Theorem 5 in [5] together with lower bound on J, much as in the proof of Lemma 8, we see that there is a small positive number μ such that either $u_p(s, t+\mu) = Ds$ or $u_p(s, t-\mu) = Ds$. Without loss of generality suppose the former, and that μ is the least possible such value. Write $\mu = r\sqrt{s}$. Now rescale to define

$$u_\flat(x_1,x_2) = s^{-1} u_p(sx_1, r\sqrt{s}x_2 + t). \tag{40}$$

Then $u_\flat$ satisfies the hypotheses on the function v of Lemma 9 (and we can suppose R is as large as we please, since we are only concerned with small s) We write $\xi_1^\flat, \xi_2^\flat$ for the derivatives of $u_\flat$. We see that from Lemma 9 that points with $\xi_1^\flat < -\zeta_1$ and $\zeta_2 < \xi_2^\flat < 2\zeta_2$ lie in a fixed bounded set.

Now write $V_\flat$ for the vector field associated to $u_\flat$, as in Section 4. Calculating the transformation under rescaling (40) we find that $V_\flat$ is *bounded*. Let ϕ be the Legendre transform of $u_\flat$ and consider the rectangle

$$Q = \{(a_1, a_2) : -(\zeta_1 + 1) \leq a_1 \leq a_2, \zeta_2 \leq a_2 \leq 2\zeta_2\}.$$

The bound on $V_\flat$ means that the determinant of the Hessian of ϕ varies by a bounded factor over Q. Since this derivative maps Q into a bounded set we get an upper bound on this determinant at each point of Q. Further, for any given ρ we get an upper bound on the Hessian over the whole ball $|\underline{a}| \leq \rho$. Now the choice of scaling, and the bound on $D(p,)$, gives bounds on the derivative of $u_\flat$ over a disc of radius 1/4, say, centred at $(1, 0)$. Since the determinat of the Hessian of ϕ is the inverse of $\det(u_\flat)_{ij}$, at the corresponding point, we obtain a *lower* bound on $\det(u_\flat)_{ij}$ over this disc. But we also have an upper bound on this determinant, by Lemma 14 of [5]. Then we deduce, just as in [4], bounds on all derivatives of $u_\flat$ on a small neighbourhood of the point $(1, 0)$ In particular

$$\left|\frac{\partial^2 u_\flat}{\partial x_2^2}\right| \leq c$$

and

$$\det((u_\flat)_{ij}) \geq c^{-1}$$

say. Now we have the transformation relations, from (40),

$$(u_\flat)_{22} = r^2 u_{22} \ , \ \det((u_\flat)_{ij}) = r^2 s \det(u_{ij}).$$

Since $s \det(u_{ij}) \leq C$ by Proposition 9, we deduce that

$$u_{22} \leq c^2 C$$

as required.

5 The vertices

Let us again take stock of our progress. We are considering a convergent sequence of data sets $(P^{(\alpha)}, A^{(\alpha)}, \sigma^{(\alpha)})$ with solutions $u^{(\alpha)}$ normalised at the centre of mass of $P^{(\alpha)}$. Our original hypothesis is that the integrals of $u^{(\alpha)}$ over $\partial P^{(\alpha)}$ are bounded, and we showed in Section 2 that the $u^{(\alpha)}$ satisfy an L^∞ bound. Then we saw in Sections 3 and 4 that the $u^{(\alpha)}$ converge away from the vertices. Our task in this section is to show that the solutions converge in neighbourhoods of the vertices. We can fix attention on a single vertex and we choose coordinates so that this vertex is the origin, that $P = P^{(\alpha)}$ is equal to the

quarter plane $\{x_1, x_2 > 0\}$ near the vertex and the measures on the two edges $\{x_1 = 0\}, \{x_2 = 0\}$ are standard. As before we usually omit the index α. For sufficiently small positive t we write

$$E(t) = t^{-1}(u(2t, 0) + u(0, 2t) - 2u(t, t)). \tag{41}$$

Our strategy is to prove

Proposition 11 *There is a bound $E(t) \leq E_0$ for all t, α.*

Of course, this is only of interest for small values of t. Given this, it is not very difficult to deduce the desired convergence around the vertex, see subsection 5.5.

Our proof of Proposition 11 is complicated, so we will first give some discussion to motivate the constructions. The bound is similar in character to the bound on the quantity D which we obtained in Section 3, and some of the same difficulties emerge in the proof. For each α choose a value t_0 which maximises the function E and set $E_{\max} = E(t_0)$. Define a function $u_\flat = u_\flat^{(\alpha)}$ by

$$u_\flat(x_1, x_2) = E_{\max}^{-1} t_0^{-1} \left(u(t_0 x_1, t_0 x_2) + \pi(x_1, x_2)\right)$$

where π is the affine-linear function chosen so that $u_\flat$ is normalised at the point $(1, 1)$. We suppose that, in the sequence (α), the maxima $E_{\max} = E_{\max}^{(\alpha)}$ tend to infinity and seek a contradiction. It is not hard to show that the $u_\flat$ converge to a convex function $u_\flat^{(\infty)}$ but the main difficulty is to rule out the possibility that

$$u_\flat^{(\infty)}(x_1, x_2) = \frac{1}{2}|x_1 - x_2|.$$

Compare with the discussion in (3.4) above, for the quantity D. To get around this we consider also the determinant function $J = \det u_{ij}$ and make various arguments with this. A crucial point is that, using the L^∞ bound from Section 2, we obtain sharp upper and lower bounds on J in terms of the Legendre transform coordinates ξ_i (Proposition 12 below). Then we consider a "perturbation" of the function E and maximise this to obtain, ultimately, the desired contradiction. (In fact we do not explicitly pass to the limit $u_\flat^{(\infty)}$ in our actual proof, making all our arguments with the smooth functions $u^{(\alpha)}$, but the reader may find it helpful to have this in mind when following the arguments.)

5.1 Volume bound

We continue with the same notation reviewed above, focussing on a vertex $(0, 0)$ and, given u, we set $\xi_i = \frac{\partial u}{\partial x_i}$. We write $J = \det u_{ij}$. Notice that for the flat model we have

$$J = (x_1 x_2)^{-1} = e^{\xi_1 + \xi_2} \tag{42}$$

Proposition 12 *There is a constant B such that*

$$B^{-1}e^{\xi_1+\xi_2} \leq J \leq Be^{\xi_1+\xi_2}$$

in a fixed neighbourhood of the vertex.

Fix some standard reference sympletic potential function u_0 (so really we have a convergent sequence $u_0^{(\alpha)}$). Let ϕ, ϕ_0 be the Legendre transforms of u, u_0 respectively. We have an elementary identity

$$\|\phi - \phi_0\|_{L^\infty} = \|u - u_0\|_{L^\infty}.$$

Clearly the u_0 are bounded and so by Theorem 2 the difference $\phi - \phi_0$ is bounded. Now take complex coordinates z_1, z_2 and set $\xi_i = \log |z_i|$, so we regard ϕ, ϕ_0 as functions of the z_i. Fix a neighbourhood N of the vertex in $\overline{P}$. Under the Legendre transform, this corresponds to some neighbourhood U of the origin in $\mathbf{C}^2$. The results of the previous section give upper and lower bounds on the difference $\log J - (\xi_1 + \xi_2)$ over the boundary of U. Since the origin is a vertex of the polygon, these functions extend to smooth functions on $\mathbf{C}^2$. The function ϕ_0 satisfies some fixed bound on the unit ball $B^4 \subset \mathbf{C}^2$ so, by the above, ϕ does also.

The results of the previous sections give us C^∞ bounds on ϕ over compact subsets of the punctured ball $B^4 \setminus \{0\}$. Let V be the volume element of the metric in these complex co-ordinates, that is $V = \det(\frac{\partial^2\phi}{\partial z_i \partial \overline{z}_j})$. So we have an upper and lower bounds on V away from the origin in B^4. The prescribed scalar curvature equation is

$$\Delta \log V = A,$$

where A is thought of as a function on $\mathbf{C}^2$ via the Legendre transform and Δ is the usual Laplace operator of the Kahler metric. Thus $|\Delta \log V| \leq C$ say. Since $\Delta\phi = 2$ we have

$$\Delta(\log V + \frac{C}{2}\phi) \geq 0 \quad , \quad \Delta(\log V - \frac{C}{2}\phi) \leq 0.$$

Thus, by the maximum principle and our bound on ϕ, the function $\log V$ over the entire ball is controlled by its values on the boundary, so we have upper and lower bounds on V over B^4. Now the chain rule gives

$$\det u_{ij} = V^{-1} \exp(\xi_1 + \xi_2)$$

and our result follows.

Next we have a simple lower bound on the determinant $\det u_{ij}$.

Lemma 10 *There is a constant $c > 0$, depending only on B above , such that $J \geq c(x_1 + x_2)^{-2}$.*

To see we argue in the same manner as in Lemma 3. We consider a point $\underline{p} = (p_1, p_2)$ in the quadrant $\{x_1, x_2 > 0\}$ and let u_p be the function obtained from u by normalising at p. Let Q be the square consisting of points (ξ_1, ξ_2) with $|\xi_i + 1| \leq 1/10$ (say) and let S be the set of points (x_1, x_2) at which the derivative of u lies in Q. The previous result implies that over S J differs by a bounded factor from $J(p)$. So we have

$$\text{Area } (S) = \int_Q J^{-1} d\xi_1 d\xi_2 \geq cJ(p)^{-1}. \tag{43}$$

Let $\underline{y}$ be a point of S and π be the affine-linear function defining the supporting hyperplane at $\underline{y}$. The zero set of π is a line which separates $\underline{y}$ and p and it follows from this that $\underline{y}$ lies in the triangle with vertices $(0,0), (p_1 + \frac{11}{9}p_2, 0), (0, p_2 + \frac{11}{9}p_1$. So the area of S is not more than $(p_1 + p_2)^2$. Rearranging (43) then gives the result.

Notice that, comparing with (42), the bound in Lemma 10 is in a sense sharp when p_1, p_2 are approximately equal.

5.2 Proof on the diagonal

Recall the definition of $E(t)$ in (41). In this subsection, and the next two, we find an *a priori* upper bound on $E(t)$.

Proposition 13 *There is a bound $E(t) \leq E_0$ for all t, α.*

Notice that $E(t)$ is not changed if we add an affine-linear function to u and that $E(t)$ is preserved by the rescaling

$$\tilde{u}(x_1, x_2) = \lambda^{-1} u(\lambda x_1, \lambda x_2).$$

Under this rescaling the function A transforms to λA. Making this rescaling, with small λ, and changing notation in the obvious way, we can suppose that u is defined on a large region in the quarter-plane $\{x_i > 0\}$. It seems simplest to take this rescaling as understood, without bringing in explicit notation. By the scaling behaviour, we can suppose that $\|A\|_{L^\infty}$ is as small as we please: let us suppose it is less than 1. It will often be convenient to work in the co-ordinates

$$t = \frac{1}{2}(x_1 + x_2), s = \frac{1}{2}(x_1 - x_2).$$

Recall that we set $E_{\max} = \max_{t>0} E(t)$. We also write $J(t)$ for the determinant $\det(u_{ij})$ evaluated at (t,t) and we write $u(t)$ for the function of one variable $u(t,t)$. For integers $n \geq 0$ let

$$\delta_n = u'(2^{-n+1}) - u'(2^{-n}).$$

Proposition 14 *There is a constant c such that then $\delta_n \leq 2 \log E_{\max} + c$ for all n,*

To prove this we observe that $E(1)$ controls the variation in the partial derivative $\frac{\partial u}{\partial s}$ over an interval in the line $t = 1$. Then we can use Lemma 14 in [5], much as in Lemma 8, to get

$$\sqrt{J(1)} \leq c \ \max((u'(2) - u'(\frac{1}{2})), E_{\max}). \tag{44}$$

Now $J(1) \geq B^{-2}J(2)\exp(u'(2) - u'(1)) = B^{-2}J(2)\exp(\delta_0)$ by Proposition 12 and so, using our lower bound of Lemma 10, $J(1) \geq c\exp(\delta_0)$. If $\exp(\delta_0)$ is large compared with $E^2_{\max}$ we must have $\sqrt{J(1)} \leq c(u'(2) - u'(\frac{1}{2}) = c(\delta_0 + \delta_1)$, so we get

$$e^{\delta_0} \leq c(\delta_0 + \delta_1). \tag{45}$$

Thus $\delta_1 \geq f(\delta_0)$ where f is the function

$$f(\delta) = c^{-1}e^{\delta} - \delta.$$

We can obviously choose a $\underline{\delta} > 1$ such that if $\delta \geq \underline{\delta}$ we have

$$f(\delta) \geq \delta^2 \geq \delta \geq \underline{\delta}.$$

Then if $\delta_0 \geq \underline{\delta}$ we have $\delta_1 \geq \delta_0^2 \geq \delta_0$. Now the whole set-up is invariant under rescaling by a power of 2, so we also have $\delta_{n+1} \geq \delta_n^2$. hence $\delta_n \geq \delta_0^{2^n}$. But by an easy argument this would imply that $u(t)$ is unbounded as $t \to 0$, contrary to what we know. So we deduce that in fact either $\delta_0 \leq \underline{\delta}$ or $\exp(\delta_0) \leq cE^2_{\max}$. Now the statement for all n follows by rescaling.

Now set

$$\Delta(t) = t^2 \max\{J(x_1, x_2) : x_1 + x_2 = 2t; |x_1 - x_2| \leq t/10\}. \tag{46}$$

Note that the factor t^2 in the definition makes this invariant under rescaling.

We introduce a parameter $\epsilon \in (0, 1)$, to be fixed later, and consider the function

$$F_\epsilon(t) = E(t) + \epsilon\Delta(t). \tag{47}$$

After scaling we can suppose this achieves its maximal value $F_{\max}$ at $t = 1$, we write $E = E(1), \Delta = \Delta(1)$. Now using the bound from Proposition 13, and (44) we get $\sqrt{\Delta} \leq c\max(E, \log E_{\max})$ so

$$E_{\max} \leq F_{\max} \leq E + (E + \log E_{\max})^2.$$

This gives

$$E_{\max} \leq c(E^2 + (\log E_{\max})^2).$$

Thus

$$E_{\max} \leq cE^2. \tag{48}$$

We can suppose that E is large (for otherwise $E_{\max}$ is not too large) then we get

$$\delta_n \leq 4\log E. \tag{49}$$

We normalise u, under the addition of affine-linear functions, at the point $(1,1)$. Then summing the δ_n, using the bound (49) and integrating the resulting bound on the $\frac{\partial u}{\partial t}$ we see that the variation of u over compact subsets of the diagonal $\{s=0\}$ is $O(\log E)$, which is small compared with the variation across the orthogonal line $\{t=1\}$, since the latter is at least E, by definition. More generally we have

Lemma 11 *For any $t_2>1$ and σ with $|\sigma|\leq 1/2$ the variation of u on the intersection of the line $\{x_1-x_2=2\sigma\}$ with the triangle $\{x_1+x_2\leq 2t_2\}$ is bounded by $c\log E$, where c depending only on t_2.*

We know that u is $O(\log E)$ on the diagonal and it follows from the definitions that u is $O(E_{\max})$ on the triangle $\{x_1+x_2\leq 3t_2\}$, say. This means that the size of the derivative of u is $O(E_{\max}d^{-1})$ where d is the distance to the boundary. Then by applying Theorem 5 of [5] we deduce that

$$J\leq cE_{\max}^2 d^{-4}. \tag{50}$$

Consider the line $\{x_1-x_2=2\sigma\}$, where we can suppose $\sigma\geq 0$, and parametrise this line by $x_1=2\sigma+\tau, x_2=\tau$. By applying Proposition 12 and the lower bound of Lemma 10 we see that

$$|\frac{\partial u}{\partial \tau}|\leq c\log(cE^4\tau^{-4}),$$

where we have used (48) to replace $E_{\max}$ by E. Integrating this we obtain the result.

In the next two subsections we prove the following two propositions.

Proposition 15 *There is a k_0, independent of ϵ, and a function $\mu(\epsilon)$ such that if at a interior maximum point for F we have $\Delta\geq k_0M$ then $E_{\max}\leq\mu(\epsilon)$.*

Proposition 16 *For any k there is an $\epsilon(k)$ and $\nu(k,\epsilon)$ such that if $\epsilon\leq\epsilon(k)$ and if at an interior maximum point for F we have $\Delta\leq kM$ then $E_{\max}\leq\nu(k,\epsilon)$.*

These two propositions complete the proof of Proposition 11. For we fix $\epsilon=\epsilon(k_0)$ and then at an interior maximum we have

$$E_{\max}\leq\max(\mu(\epsilon(k)),\nu(k_0,\epsilon(k_0))).$$

We will use a simple principle in the proofs of both of these Propositions. Write ξ_s,ξ_t for the partial derivatives of u with respect to the variables s,t. Given a point p and real numbers α,β_1,β_2 with $\beta_1<\beta_2$, let $S=S(p;\alpha,\beta_1,\beta_2)$ be the set of points (x_1,x_2) where

$$\beta_1\leq\xi_s(\underline{x})\leq\beta_2\ ,\ \xi_t\leq\alpha+\xi_t(p). \tag{51}$$

Lemma 12 *We have*

$$B^{-2}(\beta_2-\beta_1)J(p)^{-1} \leq \text{Area } (S) \leq B^2(\beta_2-\beta_1)J(p)^{-1}.$$

For the area of S is

$$\text{Area } (S) = \int_\Pi J^{-1} d\xi_s d\xi_t, \tag{52}$$

where Π is the region in the (ξ_s, ξ_t) plane defined by the inequalities (51), and we have abused notation by regarding J as a function of ξ_s, ξ_t in the obvious way. Now the volume bound of Proposition 12 gives

$$B^{-2}J(p)e^{\xi_t(p)-\xi_t} \leq J(\xi_s,\xi_t) \leq B^2 J(p) e^{\xi_t(p)-\xi_t}, \tag{53}$$

and the result follows by integrating the exponential function over Π.

5.3 Proof of Proposition 14

We fix values t_0, t_1, t_2, say for definiteness $t_0 = 1/5, t_1 = 1/4$ and $t_2 = 2$. Let R be the rectangle $\{|s| \leq 1/10, t_0 \leq t \leq t_2\}$.

Recall that the definition of Δ involves maximising over an interval $|s| \leq t/20$. Suppose that the maximum is achieved at a point p, where $t = 1$ and $s = s_0$. (Of course, we can suppose $t = 1$ by rescaling.) So $|s_0| \leq 1/20$ and p lies inside R.

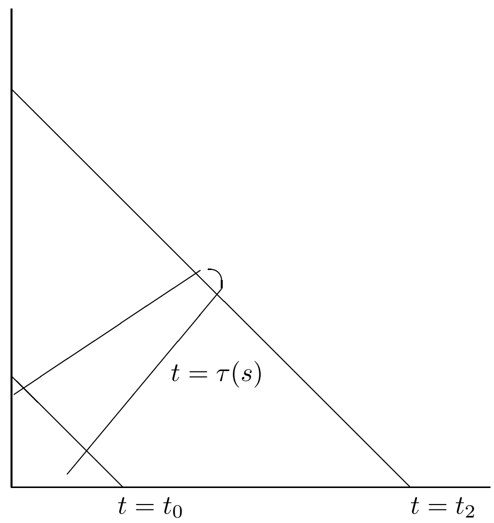
Now the proof proceeds by the following steps.

Step 1 *Claim: There is a* c_1 *such that* $|\frac{\partial u}{\partial s}| \leq c_1 E$ *on* R.

For on the line segment $\{t = 1, |s| \leq 1\}$ we have a bound $|u| \leq E|s|$. Using Lemma 11, this gives an $O(E)$ bound on u over the interior region $|s| \leq 1, t \leq 2$. Since R lies within the interior of this set, convexity gives an $O(E)$ bound on the derivative over R.

Now we consider the set $S = S(p; \alpha, -c_1E, c_1E)$, with c_1 as above. The curve $\{\frac{\partial u}{\partial t} = \alpha\}$ is the graph of a function $t = \tau(s)$. By item (1) above the intersection $S \cap R$ is just the set defined by the three conditions

$$-1/10 \leq s \leq 1/10, t_0 \leq t \leq t_2, t \leq \tau(s).$$



Step 2 *Claim: We can choose $\alpha > 0$, depending only on B, so that for any point q on the graph $t = \tau(s)$ and any point p' with $s = s_0, t \leq 1$ we have*

$$J(q) < \frac{t_2 - t_1}{1 - t_1} J(p').$$

For, since the partial derivative $\frac{\partial u}{\partial t}$ is monotone on the line $s = s_0$ we have

$$J(p') \geq B^{-1} J(p),$$

whereas, by the inequality of Proposition 12,

$$J(q) \leq B e^{-\alpha} J(p).$$

So we just need to choose $\alpha > 0$ and bigger than $\log\left(B^2 \frac{t_2 - t_1}{1 - t_1}\right)$.

Now we fix α as above. By Lemma 12, the area of S is at most $cE/J(p) = cE/\Delta$. So, by choosing k_0 large (as allowed in the statement of Proposition 14), we can suppose the area of S is as small as we please. Fix a suitably small number δ— for definiteness we can take $\delta = 1/100$—and choose k_0 so that the area of S is less than $\delta(t_1 - t_0)$.

Note that, since $\alpha > 0$, we have $\tau(s_0) > 1$, by monotonicity of the partial derivative.

Step 3 *Claim: There are s_-, s_+ with $|s_\pm - s_0| \leq \delta$ and $s_- < s_0 < s_+$ such that $\tau(s_\pm) \leq t_1$.*

If there is no such s_+ then S contains the rectangle $t_0 \leq t \leq t_1, s_0 \leq s \leq s_0 + \delta$. (Notice that our choices imply that this rectangle lies inside R.) But

this contradicts the fact that the area of S is less than $\delta(t_1 - t_0)$. Similarly for s_-.

To sum up so far we have shown that the set S must contain a very thin "finger", extending out from the region $\{t \leq t_1\}$ and containing the point p where $t = 1$.

Let s_+ be the least among the values satisfying the conditions of the claim above and s_- be the largest. Then $\tau(s_\pm) = t_1$ and $\tau > t_1$ on the open interval (s_-, s_+).

Let Ω be the set where $s_- \leq s \leq s_+$, $t_1 \leq t \leq t_2$ and $t \leq \tau(s)$. For $t' \in [t_1, t_2]$ let $I_{t'}$ be the intersection of Ω with the line $t = t'$ and let $j(t')$ be the maximum of J over $I_{t'}$. Thus $j(1) \geq \Delta$. Let $G \subset [t_1, t_2]$ be the set of values t^* such that for all $t' \in (t_1, t^*]$ the maximum $j(t')$ is attained at an interior point of $I_{t'}$ (i.e. not at points in the graph of τ).

Step 4 *Claim:* 1 *is contained in* G*.*

This follows from the Claim in Step 2, since for $t' \leq 1$ the point p' with co-ordinates $t = t', s = s_0$ lies in $I_{t'}$ and $J(p')$ is strictly less than the value of J at any point on the graph.

Now the crucial idea in the proof is to show that this thin "finger" must actually extend to meet the line $\{t = t_2\}$.

Step 5 *Claim: If* $t^* \geq 1$ *and* $t^* \in G$ *then* j^{-1} *is a concave function on the interval* $[t_1, t^*]$*.*

This is similar to the proof of Proposition 9. The function $F = J^{-1}$ satisfies the linear equation $u_{ij}F^{ij} = -A$ and $A \geq 0$. Then the assertion follows from the maximum principle applied to $F - ct$ for suitable values of c.

Step 6 *Claim:* t_2 *is in* G*.*

This follows from a continuity argument. From its definition, G is open. So long as t^* lies in G we have

$$j(t^*)^{-1} \leq \frac{t^* - t_1}{1 - t_1} j(1)^{-1} \leq \frac{t^* - t_1}{1 - t_1} J(p)^{-1},$$

by convexity. Suppose $1 \leq t^* \leq t_2$. Recall that we arranged that for any point q on the graph of τ

$$J(q)^{-1} > \frac{t_2 - t_1}{1 - t_1} J(p)^{-1}.$$

The strict inequality implies that G is closed.

Step 7 *Claim: There is a point p'' with co-ordinates $s'' \in (s_-, s_+)$ and $t = t_2$ such that $J(p'') > \frac{1-t_1}{t_2-t_1} J(p)$*

This follows from the concavity of j^{-1}, as above.

Now by the choice of δ we have $|s''| \leq t_2/20$ so the point above is one of those considered in the definition of $\Delta(t_2)$ and we have

$$\Delta(t_2) \geq \frac{1-t_1}{t_2-t_1} t_2^2 \Delta.$$

Now with the definite choices of t_i made above this inequality is $\Delta(t_2) \geq (1+\sigma)\Delta$ with $\sigma = 5/7 > 0$.

We can now complete the proof. From Lemma 11 we know that $u(0,0)$ is $O(\log E)$ and the convexity of u on the boundary implies that

$$u(2t_2, 0) \geq t_2 u(1,0) - c \log E, u(0, 2t_2) \geq u(0,1) - c \log E.$$

Also $u(t_2, t_2)$ is $O(\log E)$, again by Lemma 11, so from the definition of $E(t)$ we have

$$E(t_2) \geq E - c \log E.$$

So

$$F_\epsilon(t_2) \geq E + \epsilon(1+\sigma)\Delta - c \log E \geq F_\epsilon(1) + \epsilon\sigma k_0 E - c \log E.$$

Proposition 14 follows from (48) and the fact that $F_\epsilon(t_2) \leq F_\epsilon(1)$.

5.4 Proof of Proposition 15

Recall from the statement of the Proposition that we are supposing that $\Delta \leq kE$. The main idea in the proof will be in part complementary to that of Proposition 14, in that we invoke a *lower* bound on the area of a suitable set S. As before we suppose that the maximum of F_ϵ is attained at $t = 1$, and let p be the point on the line $\{t = 1\}$ where the maximum in the definition of Δ is achieved. Our argument again employs certain parameter values t_0, t_2 for the t-coordinate, but this time we will choose $t_0 < 1$ very small, so that

$$\frac{1}{2} t_0^2 \leq \frac{1}{100 B^2 k}, \tag{54}$$

and $t_2 > 1$ very large, so that

$$t_2^2 \geq 100 B^4 k. \tag{55}$$

For any t we have

$$E(t) \leq F_\epsilon(t) \leq F_\epsilon(1) = E + \epsilon\Delta \leq (1 + \epsilon k)E.$$

Write $U(t) = u(2t, 0) + u(0, 2t)$, so $E = U(1)$. Over the fixed range, $t \leq t_2$, our bound in Lemma 11, on the diagonal, gives $U(t) \leq tE(t) + c \log E$. Hence

$$U(t) \leq t(1 + \epsilon k)E + c \log E = t(1 + \epsilon k)U(1) + c \log E \tag{56}$$

When ϵ is small, convexity of the function U forces $E^{-1}U(t)$ to be close to the linear function t (assuming of course that E is large), over the range $t \leq 1$. Further, each summand $u(2t, 0), u(0, 2t)$ is positive and convex and this forces

$$u(2t, 0) = tu(2, 0) + O(\log E + \epsilon E) \quad , \quad u(0, 2t) = tu(0, 2) + O(\log E + \epsilon E).$$

To express this differently, write $u(2, 0) = \lambda_1 E, u(0, 2) = \lambda_2 E$, so $\lambda_i \geq 0$ and $\lambda_1 + \lambda_2 = 1$. Define

$$V(x_1, x_2) = \frac{1}{2} \max\left(\lambda_1(x_1 - x_2), \lambda_2(x_2 - x_1)\right).$$

Then on compact subsets of the boundary of the quarter-plane the function $E^{-1}u$ differs from V by $O(E^{-1} \log E + \epsilon)$. But now Lemma 11 implies that $E^{-1}u$ differs from V by $O(E^{-1} \log E + \epsilon)$ over the whole region $\{t \leq t_2\}$ in the quarter plane. Set

$$\beta_1 = \left(\frac{1}{2}(\lambda_1 - \lambda_2) - \frac{1}{4}\right) E \quad , \quad \beta_2 = \left(\frac{1}{2}(\lambda_1 - \lambda_2) + \frac{1}{4}\right) E.$$

(The reader may find it easiest to think first of the symmetrical case when $\lambda_1 = \lambda_2 = 1/2$.) We consider the set $S = S(p; \alpha, \beta_1, \beta_2)$, where $\alpha > 0$ and

$$e^{\alpha} \geq 10B^2 k. \tag{57}$$

(The reason for these choices will emerge presently.) The crucial observation is

Lemma 13 *Thee intersection of S with the set $\{t_0 \leq t \leq t_2\}$ is contained in a strip $|s| \leq \eta$ where $\eta = c\epsilon$, once E is sufficiently large.*

The proof is straightforward, using the preceding discussion. (Only the constraint $\beta_1 \leq \xi_s \leq \beta_2$ is relevant here: the statement is valid for any α.)

Now we proceed with the following steps.

Step 1 *Claim: If E is sufficiently large, the variation of $\frac{\partial u}{\partial s}$ on the line $t = 1$ across the strip $|s| \leq \eta$ is at least $E/2$.*

We know that $\frac{\partial u}{\partial s}$ varies from $-\infty$ to ∞ across the whole interval $t = 1, |s| \leq 1$ and by the Lemma 13 the points where values β_1, β_2 are attained must lie in this strip. Then the claim follows from the fact that $\beta_2 - \beta_1 = E/2$.

Step 2 *Claim: There is a $k' \geq 1/10$, such that, if E is sufficiently large, there is a point p' in the segment $|s| \leq \eta, t = 1$ with $\beta_1 \leq \xi_s(p') \leq \beta_2$ and $J(p') \geq k'E$.*

To see this we use an integral identity just as in [5], Lemma 17. This gives a formula, in terms of A, for the integral over the line segment $t=1, |s|\leq 1$ of u^{tt} (in an obvious notation). Using the formula for the inverse of a 2×2 matrix, we can write this is

$$\int J^{-1} d\xi_s,$$

where $\xi_s = \frac{\partial u}{\partial s}$ is regarded as a parameter on the line segment $|s|\leq 1, t=1$. From this formula one sees that the integral is at most 5, when $|A|_{L^\infty}\leq 1$, as we are supposing. In particular the same integral over the sub-segment $|s|\leq \eta, t=1$ is bounded above by $k'/2 \leq 5$ say. Since the variation in ξ_s over this subsegment at least $E/2$ there must be a point where $J\geq k'E$, with $k'\geq 1/10$.

Step 3 *Claim: If ϵ is sufficiently small and E is sufficiently large then $\Delta \geq k'E$.* We just choose ϵ so that $\eta < 1/20$ and the point p' is one of those considered in the definition of Δ.

Step 4 *Claim: If ϵ is sufficiently small, and E sufficiently large, the set S intersects the line segment $\{t=t_2, |s|\leq \eta\}$*

First, since $\frac{J(p')}{J(p)} \geq (10k)^{-1}$ it follows from the choice of α and Proposition 12 that p' lies in S. By Lemma 12, the area of S is at least $\frac{E}{2\Delta B^2}\geq (2kB^2)^{-1}$ and by the choice of t_0 this is more than twice the area of the triangle $\{t\leq t_0\}$. Suppose S does not intersect the line segment as claimed. Since p' lies in S and S is connected it follows from Lemma 13 that S is contained in the union of the triangle $\{t\leq t_0\}$ and the strip $\{t_0\leq t\leq t_2, |s|\leq \eta\}$. So the area of this strip must be at least half the area of S. But the area of strip is $2(t_2-t_0)\eta = 2c(t_2-t_0)\epsilon$, so this is impossible when ϵ is sufficiently small.

Step 5 *Claim:* $\Delta(t_2)\geq 2\Delta$.

Consider the point p'' whose existence is established in the previous step. Using Proposition 12, $J(p'')\geq B^2 e^{-\alpha} J(p) = B^2 e^{\alpha}\Delta$. We can assume that η is small, so p'' is a point considered in the definition of $\Delta(t_2)$ and $\Delta(t_2)\geq t_2^2 J(p'')$. Now the claim follows from the choice of t_2.

Now we can complete the proof. By convexity of the function U and fact that $U(0)$ is $O(\log E)$ we have

$$U(t_2)\geq t_2 E - c\log E.$$

From the bound on $u(t_2,t_2)$ we deduce that

$$E(t_2)\geq E - c\log E.$$

So

$$F_\epsilon(t_2) \geq E + \epsilon\Delta(t_2) - c\log E \geq E + 2\epsilon\Delta - c\log E.$$

Now the fact that $F_\epsilon(t_2) \leq F_\epsilon(1)$ gives

$$E + 2\epsilon\Delta - c\log E \leq E + \epsilon\Delta,$$

so $\epsilon\Delta \leq c\log E$. Now by Step 3, $\Delta \geq k'E$ so

$$\epsilon k'E \leq c\log E,$$

which gives the required bound on E.

5.5 Completion of proof of Main Theorem

We need to control a solution $u^{(\alpha)}$ in a neighbourhood of a vertex, and we can take standard co-ordinates around the vertex as in the previous section. For each point $p = (p_1, p_2)$ we define $D_1(p)$ as in (3.2)

$$D_1(p) = \frac{1}{p_1}(u(0, p_2) - u(p_1, p_2) - \frac{\partial u}{\partial x_1}(p_1, p_2).$$

Of course we have a similar quantity $D_2(p)$ defined by interchanging the co-ordinates. We prove

Proposition 17 *There is a a priori bound $D_1(p), D_2(p) \leq D$, valid for all solutions $u = u^{(\alpha)}$ in our sequence, and all points p near to a vertex.*

Given this it is straightforward to adapt the proofs of proposition 6 to deduce Theorem 1, arguing just as in (4.1).

When $p_1 = p_2$ a bound on $D_i(p)$ follows immediately from what we have proved in the previous section. More generally, the only new issues arise when p_1 is much less than p_2. We adapt the argument of Section 3. We choose a point where $D_1(p)$ is maximal and by rescaling we can suppose that $p_2 = 1$. If u is normalised at $(1, 1)$, as in the previous subsection, then we have *a priori* L^∞ bounds on u over compact subsets. The only difficulty in applying the argument of Section 3 would occur if the "scale" λ is not small. Then the limit $u_\flat$ would still be defined only on a quarter plane and the problem would come when, taking the limit in seeking a contradiction, the $u_\flat$ converge to an affine-linear function on the boundary. To rule this out we need an *a priori* "strict convexity" bound on the restriction of u to the x_2-axis. Thus the crucial thing is to prove

Proposition 18 *There are $r > 1$ and $\eta > 0$ such that*

$$\frac{\partial u}{\partial x_2}(0, r) - \frac{\partial u}{\partial x_2}(0, 1) \geq \eta,$$

for all functions u obtained by rescaling a $u^{(\alpha)}$

Given this proposition it is very to adapt the arguments of Section 3 to prove Proposition 16, on the lines indicated above.

To prove Proposition 18, we begin by considering the derivative $\xi_t = \frac{\partial u}{\partial t}$ on the diagonal. We have upper and lower bounds on the determinant J at the point $(1,1)$ and, under rescaling, these give

$$c' t^{-2} \geq J(t,t) \geq c t^{-2}.$$

Then Proposition 12 gives

$$\xi_t \geq 2\log t - c$$

for $t \geq 1$. Now write $\frac{\partial u}{\partial x_2} = \xi_t - \xi_s$ so

$$\frac{\partial u}{\partial x_2}(t,t) \geq 2\log t + \xi_s(t,t) - c.$$

The L^∞ bound on u implies a bound on $\frac{\partial u}{\partial x_2}(1,1) - \frac{\partial u}{\partial x_2}(0,1)$. By considering the rescaling behaviour we get a fixed bound on $\frac{\partial u}{\partial x_2}(t,t) - \frac{\partial u}{\partial x_2}(0,t)$ for all t. So

$$\frac{\partial u}{\partial x_2}(0,t) \geq 2\log t + |\xi_s(t,t)| - c. \tag{58}$$

Thus it suffices to show that $|\xi_s(t,t)|$ is small, for large t, compared with $2\log t$. To this end we first define

$$T = T(u) = |\xi_s(2,2) - \xi_s(1,1)|.$$

(In fact u has been normalised so that $\xi_s(1,1) = 0$, but it clearer to write the definition this way.) By our L^∞ bounds we have $T \leq C_0$ say.

For integer $\mu > 1$ let Ω_μ be the region $\{2^{-\mu} \leq t \leq 4\}$ in the quarter-plane. Let $E_\mu(u)$ be the integral of the quantity $|F|^2$ (as defined in [4], [5]) over Ω_μ. (This is essentially the square of the L^2 norm of the Riemann curvature tensor over the corresponding piece of a 4-manifold.) Given a positive number C, let $\mathcal{A}_C$ be the set of convex functions u on the closed triangle $\{t \leq 4\}$, normalised at $(1,1)$ such that

1. u satisfies equation (1), with $\|A\|_{C^2} \leq C$;
2. $\|u\|_{L^\infty} \leq C$;
3. $\det(u_{ij}) \geq C^{-1}$ everywhere;
4. $\|V\|_{L^\infty} \leq C$, where V is the vector field associated to u;
5. u satisfies Guillemin boundary conditions, with the standard measure, along the x_i-axes;
6. $\det(u_{ij}) \geq C^{-1} t^{-2}$ at the point $(\frac{t}{2}, \frac{t}{2})$.

Then we have

Lemma 14 *For any $C, \epsilon > 0$ there is an integer μ and a $\delta > 0$ such that if $u \in \mathcal{A}_C$ and $E_\mu \leq \delta$ we have $T(u) \leq \epsilon$.*

For our application we fix $\epsilon < 2 \log 2$.

Assuming this lemma for the moment, we complete the proof of Proposition 17. For integers $n \geq 1$ set

$$T_n = |\xi_s(2^n, 2^n) - \xi_s(2^{n-1}, 2^{n-1})|.$$

Thus, rescaling by a factor 2^n, we can apply our bounds on T to give bounds on T_n. For suitable C, all of the conditions defining $\mathcal{A}_C$ hold for our $u^{(\alpha)}$, so $T_n \leq C_0$ for all n, and either $T_n \leq \epsilon$ or the integral of $|F|^2$ over the region $\{2^{n-\lambda} \leq t \leq 2^{n+1}\}$ exceeds δ. Now use the fact that we have a fixed bound on the L^2 norm of F over the whole polygon, so we can find a large integer M such that

$$\int_P |F|^2 \leq M\delta.$$

It follows that there are at most $M(\lambda + 1)$ values of n for which T_n exceeds ϵ. Thus

$$|\xi_s(2^n, 2^n) - \xi_2(1, 1)| \leq M(\lambda + 1)C_0 + n\epsilon. \tag{59}$$

Combining (59) (with $t = 2^n$) with (58) we establish Proposition 17.

It only remains to prove Lemma 14. Arguing by contradiction, we suppose we have a sequence of functions $u^{(\beta)} \in \mathcal{A}_C$ with the integral of $|F|^2$ over Ω_p tending to zero, for a sequence p tending to infinity, and the sequence of quantities $T(u^{(\beta)})$ does not tend to zero. Using the first three conditions in the definition of $\mathcal{A}_C$ and the arguments of [4] we see that, taking a subsequence, we can suppose the sequence converges in C^4 on compact subsets of the interior. The limit clearly has $F = 0$, i.e. describes a flat metric. The vector field V associated to this limit is constant and a straightforward Stokes' Theorem argument (similar to that in Proposition 5), using the boundary conditions and the fourth condition before taking the limit, shows that V_∞ is the vector field $\frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_2}$. There is a simple classification of locally-defined functions u_∞ with $F = 0$ and it is easy to read off from this that in the limit ξ_s is constant along the diagonal, which gives the desired contradiction. (The point here is that the fifth condition forces the limit to blow up at the origin.)

6 Blow-up limits

6.1 The Joyce construction

We recall a construction, due to Joyce, of explicit solutions of equation (1), with $A = 0$, that is, metrics of zero scalar curvature. The original reference is [7], but we follow the approach of Calderbank and Pedersen in [2]. An elementary

derivation of this construction (and a generalisation to other equations) from the point of view of this paper is given in the note [6].

Consider the linear PDE for a function $\xi(r, H)$, where $r > 0$,

$$\frac{\partial^2 \xi}{\partial H^2} + r^{-1}\frac{\partial}{\partial r}\left(r\frac{\partial \xi}{\partial r}\right) = 0. \tag{60}$$

This is familiar as the equation defining axi-symmetric harmonic functions in cylindrical co-ordinates on $\mathbf{R}^3$. Given a pair of solutions ξ_1, ξ_2 to (60), we set

$$P_i = \frac{\partial \xi_i}{\partial H}\ ,\ Q_i = \frac{\partial \xi_i}{\partial r}.$$

and write $\Delta = P_1Q_2 - Q_1P_2$. We assume that $\Delta > 0$ everywhere. We introduce two further angular co-ordinates θ_1, θ_2 and consider the four dimensional Riemannian metric

$$g = \frac{r\Delta}{2}(dH^2+dr^2)+\frac{r}{2\Delta}\left((P_2^2+Q_2^2)d\theta_1^2 - 2(Q_1Q_2+P_1P_2)d\theta_1 d\theta_2 + (P_1^2+Q_1^2)d\theta_2^2\right).$$

The main result is that this is a Kahler metric of zero scalar curvature. To relate this to the equation (1), we introduce another linear equation

$$\frac{\partial^2 x}{\partial H^2} + r\frac{\partial}{\partial r}\left(r^{-1}\frac{\partial x}{\partial r}\right) = 0. \tag{61}$$

Given a solution $\xi(r, h)$ to (60), the first order system

$$\frac{\partial x}{\partial r} = r\frac{\partial \xi}{\partial H}\ \ ,\ \ \frac{\partial x}{\partial H} = -r\frac{\partial \xi}{\partial r},$$

is consistent and has a solution x, unique up to a constant. Further x satisfies the equation (61). So starting with a pair of solutions ξ_1, ξ_2 to (60) we get a pair of solutions x_1, x_2 to (61), but we introduce a sign and interchange of labels so that

$$\frac{\partial x_1}{\partial r} = \frac{\partial \xi_2}{\partial H}\ \ ,\ \ \frac{\partial x_2}{\partial r} = -\frac{\partial \xi_1}{\partial H},$$

etc.. Then x_i and ξ_j give the co-ordinates we are familiar with in this paper: the equations imply that the 1-form $\sum \xi_i dx_i$ is closed, so there is, at least locally, a function u with $du = \sum \xi_i dx_i$. If we regard u as a function of (x_1, x_2) then we get a solution of (1), with $A = 0$. Conversely any solution arises in this manner away from the critical points of $\det(u_{ij})$. In fact the construction gives $r = \det(u_{ij})^{-1/2}$.

An important special case occurs when $\log J$ is an affine-linear function of ξ_1, ξ_2. In differential geometric terms, our metric is then Ricci-flat. Making an affine change of variable we may suppose that $\xi_2 = \log r$. So $P_2 = 0$ and $Q_2 = r^{-1}$. It is easy to check that the metric g is the same as that given by the well-known Gibbons-Hawking construction, using the harmonic function $\frac{\partial \xi_1}{\partial H}$ on $\mathbf{R}^3$.

With this background in place, we can move on to consider the particular metrics we are interested in. Consider first the case of flat space, so

$$u = x_1 \log x_1 + x_2 \log x_2$$

on the quadrant $\{x_1, x_2 > 0\}$. Then one finds that

$$\xi_1 = \log F_-(H, r), \xi_2 = \log F_+(H, r)$$

where

$$F_\pm(H, r) = \frac{1}{2}\left(\pm H + \sqrt{H^2 + r^2}\right).$$

The harmonic function $\log F_-$ is the potential associated to a uniform charge distribution on the half-line $r = 0, H > 0$ and $\log F_+$ to the half-line $r = 0, H < 0$. So, as we see from the formulae, $F_\pm$ has a logarithmic singularity along the corresponding half-line and

$$F_+ + F_- = 2 \log r.$$

The corresponding functions x_i are just $x_1 = F_-, x_2 = F_+$.

Now given $a_1, a_2 > 0$ set

$$\xi_1 = \log F_-(r, H) - a_2 H + 1, \xi_2 = \log F_+(H, r) + a_1 H + 1.$$

(The addition of the constant 1 makes no change to the geometry but will be convenient later.) These are obviously harmonic functions and the corresponding functions x_i are

$$x_1 = F_- + \frac{a_1 r^2}{4} \ , \ x_2 = F_+ + \frac{a_2 r^2}{4}.$$

We want to find the "symplectic potential" $u(x_1, x_2)$ which describes this solution. Set $y_1 = F_-, y_2 = F_+$ so that $y_2 - y_1 = H$ and $y_1 y_2 = r^2/4$. Thus

$$x_1 = y_1 + a_1 y_1 y_2 \ , x_2 = y_2 + a_2 y_1 y_2.$$

and

$$\xi_1 = \log y_1 + a_2(y_1 - y_2) + 1 \ , \ \xi_2 = \log y_2 + a_1(y_2 - y_1) + 1.$$

The defining condition for u is $du = \xi_1 dx_1 + \xi_2 dx_1$ which, after some cancellation using the fact that

$$dx_i = dy_i + a_i(y_1 dy_2 + y_2 dy_1) \tag{62}$$

is the differential

$$(\log y_1 dx_1 + \log y_2 dx_2) + dx_1 + dx_2 + (y_1 - y_2)(a_2 dy_1 - a_1 dy_2).$$

Set $V = x_1 \log y_1 + x_2 \log y_2$ so

$$dV = \log y_1 dx_1 + \log y_2 dx_2 + \frac{x_1}{y_1} dy_1 + \frac{x_2}{y_2} dy_2.$$

Then

$$du - dV = dx_1 + dx_2 + (1 - a_1 y_2) dy_1 + (1 - a_2 y_2) dy_2 + (y_1 - y_2)(a_2 dy_1 - a_1 dy_2).$$

Using (62), expanding and cancelling terms, one finds that

$$du - dV = a_2 y_1 dy_1 + a_1 y_2 dy_2,$$

so we can take

$$u = V + \frac{1}{2}(a_2 y_1^2 + a_1 y_2^2).$$

So far, we have been working rather formally—ignoring the precise domains of our functions—and now we return to a global point of view. One can verify that, for any $a_i > 0$, the map

$$(y_1, y_2) \mapsto (y_1 + a_1 y_1 y_2, y_2 + a_2 y_1 y_2)$$

yields a diffeomorphism from the quadrant $\{y_i \geq 0\}$ (regarded as a manifold with a corner) to itself. So we have an inverse diffeomorphism given by functions $y_i = y_i(x_1, x_2)$, which can be given by explicit formulae, as below. Now define

$$u(x_1, x_2) = x_1 \log y_1 + x_2 \log y_2 + \frac{1}{2}(a_2 y_1^2 + a_1 y_2^2). \tag{63}$$

Then u is a convex function on the quarter-plane, satisfying Guillemin boundary conditions along the axes and $u^{ij}_{ij} = 0$.

Multiplying a_1, a_2 by the same non-zero factor does not change the metric up to isometry. When $a_1 = a_2$ we have $\xi_1 + \xi_2 = \log r + 2$ and the metric is Ricci flat, given by the Gibbons-Hawking construction using the harmonic function $\frac{1}{|r|} + 1$ on $\mathbf{R}^3$. (We need to make a linear change of co-ordinates to ξ_s, ξ_t to fit in with the discussion above.) This is a standard description of the recognize the *Taub-NUT* metric on $\mathbf{R}^4$, which is well-known to be complete, with curvature in L^2. When $a_1 \neq a_2$ the metric is not Ricci-flat. It is easy to see that it is complete: the author expects, but has not yet checked in detail, that the curvature is in L^2. (The definitions above make sense when one of the a_i is zero, and we still get a metric on $\mathbf{R}^4$. But in this case the curvature is definitely *not* in L^2, so we exclude it.)

We want to discuss the asymptotic behaviour of one of these solutions for large $\underline{x} = (x_1, x_2)$. It is convenient to make a linear change of variable

$$\sigma = a_2 x_1 - a_1 x_2 \quad , \quad \tau = a_1 x_2 + a_2 x_1.$$

(So in the Taub-NUT case, when $a_1 = a_2 = 1/2$ these coincide with the co-ordinates s, t we used in Section 5.) Then we can solve for y_i to find

$$2y_1 = (\sigma - 1) + \sqrt{\sigma^2 + \frac{\tau}{a_2} + 1}, 2y_2 = (1 - \sigma) + \sqrt{\sigma^2 + \frac{\tau}{a_1} + 1}. \tag{64}$$

We consider the behaviour when τ is large in the three sectors $\sigma > \epsilon\tau, |\sigma| \leq \tau, \sigma < -\epsilon\tau$, for fixed ϵ. If $\sigma > \epsilon\tau$ we have $y_1 \sim \sigma$ and $y_2 = O(1)$ while if $\sigma < -\epsilon\tau$ we have $y_1 = O(1)$ and $y_2 \sim \sigma$. The asymptotics in a sector $-\epsilon\tau \leq \sigma < \epsilon\tau$ are more complicated, but on the line $\sigma = 0$ we have $y_i = O(\sqrt{\tau})$. Now substituting in the formula (63) for u we find that when $\sigma = 0$, $u = O(\tau \log \tau)$. This is the same growth rate as in the Euclidean case. When $\sigma > \epsilon\tau$ we have $u \sim \frac{a_1}{2}\sigma^2$ while if $\sigma < -\epsilon\tau$ we have $u \sim \frac{a_2}{2}\sigma^2$. Thus u grows much faster away from the line $\sigma = 0$ than it does along this line, but the growth rates in the two regions $\pm\sigma \geq \epsilon\tau$ are different.

6.2 Discussion

Suppose we have a convergent sequence of data sets $(P^{(\alpha)}, \sigma^{(\alpha)}, A^{(\alpha)})$, with solutions $u^{(\alpha)}$ but the limit $(P^{(\infty)}, \sigma^{(\infty)}, A^{(\infty)})$ does *not* satisfy the positivity condition on $L = L_{A^{(\infty)},\sigma^{(\infty)}}$ discussed in the Introduction. How do the solutions $u^{(\alpha)}$ behave as $\alpha \to \infty$? By the results of [3], there is an affine-linear function λ which changes sign on P, such that $L(\lambda^+) = 0$, where $\lambda^+ = \max(\lambda, 0)$. Suppose for the moment that this is the unique function with this property, up to a factor. The line (or "crease") $\{\lambda = 0\}$ divides P into two pieces. What we expect is that on the interior of each piece the $u^{(\alpha)}$ converge, after suitable normalisation, but the normalisations required are different, and if we normalise u^α on the region $\{\lambda < 0\}$ then on the other region $\{\lambda > 0\}$ the functions blow up as

$$u^\alpha \sim n_\alpha \lambda^+$$

with scalars $n_\alpha \to \infty$.

We can easily write down explicit examples of this behaviour, which is essentially a one-dimensional phenomenon. Fix a family of even functions f_ϵ on the interval $[-1, 1]$, parametrised by $\epsilon \in [0, 1]$, with $f_\epsilon(x) = x^2 + \epsilon^2$ for $|x| \leq 1/2$, with $f_\epsilon(x) = 1 - |x|$ for $|x|$ close to 1 and with $f_\epsilon(x) > 0$ except when $|x| = 1$ or $x = \epsilon = 0$. We can obviously do this in such a way the family is smooth in both variables. Set $a_\epsilon = f_\epsilon''$. Then for $\epsilon > 0$ the one-dimensional version of (1), which is

$$\left(\frac{1}{U(x)''}\right) = -a_\epsilon, \tag{65}$$

has a solution U_ϵ which is smooth in $(-1, 1)$ and satisfies the Guillemin boundary condition at ± 1. We just take U_ϵ to be the solution of the elementary equation

$$U_\epsilon'' = f_\epsilon^{-1},$$

normalised to $U_\epsilon(0) = U_\epsilon'(0) = 0$, say. Equally obviously, the family U_ϵ is unbounded on any neighbourhood of 0, as $\epsilon \to 0$, because the derivative U_ϵ' has limit $|x|^{-1}$ which is not integrable. If, on the other hand, we define $U_{+,\epsilon}$ to be the solution of (65) with $U_{+,\epsilon}(1/2) = U_{+,\epsilon}'(1/2) = 0$, then the $U_{+,\epsilon}$ converge as

$\epsilon \to 0$ on the interval $(0,1]$. Similarly there is a family $U_{-,\epsilon}$ normalised at $-1/2$ and converging over $[-1,0)$. For $\epsilon > 0$ we have

$$U_{+,\epsilon}(x) - U_{-,\epsilon}(x) = n_\epsilon x,$$

where $n_\epsilon \to \infty$ (and in fact $n_\epsilon \sim \log \epsilon^{-1}$). To get a two-dimensional example we can simply take P to be the square $[-1,1]^2 \subset \mathbf{R}^2$ and $A_\epsilon(x_1,x_2) = a_\epsilon + 1$. Then for $\epsilon > 0$ there is a solution $u_\epsilon(x_1,x_2) = U_\epsilon(x_1) + V(x_2)$, where V is the symplectic potential for the round metric on the 2-sphere. In terms of Riemannian geometry in four-dimensions described by this family: in the region corresponding to $[-1/2,1/2] \times [-1,1]$ the 4-manifolds have the form

$$H_\epsilon \times S^2$$

with the product of the round metric of curvature $+1$ on S^2 and a metric of curvature -1 on H_ϵ, approaching a pair of "cusps" as $\epsilon \to 0$. The curvature of these metrics is bounded, uniformly in ϵ, but the diameter tends to infinity and the injectivity radius to 0.

This gives a model, albeit somewhat conjectural, for the behaviour near a "crease" $\{\lambda = 0\}$ in the general case, *provided that this crease does not pass through a vertex of P*. Suppose on the other hand that we are in this situation, and take the standard model with the vertex the origin and P coinciding locally with the quarter-plane $\{x_i > 0\}$. The crease is a line $x_2 = bx_1$, where $b > 0$. What seems likely to be true is that, near the origin, the solutions $u^{(\alpha)}$ are modelled on the zero scalar curvature metric discussed above with parameters $a_1 = 1, a_2 = b$, then scaled down by a factors r_α, with $r_\alpha \to 0$ as $\alpha \to \infty$ (which is the same as taking parameters $a_1 = r_\alpha, a_2 = br_\alpha$. In other words, we expect these model solutions to appear as the "blow-up" limits. This picture is consistent with the discussion in the preceding subsection of the asymptotics of the model solutions. Notice that (if this picture is correct) then in this situation the curvature of the $u^{(\alpha)}$ is not bounded uniformly in the family, in contrast to the previous case.

Of course we can envisage somewhat more complicated situations in which we have several "creases", dividing the polygon into more parts. This is discussed in [3], and taken much further by Szelyhidi in [8]. Szekelyhidi's work also suggests very strongly that a similar picture holds for the limiting behaviour of the Calabi flow. Notice also that this (conjectural) picture is quite in line with the more general situation, in four-dimensional Riemannian geometry, described by Anderson in [1].

These models are also useful in understanding the issues involved in the existence proofs. Consider the first "product" example, but apply an affine transformation so the domain P is now a parallelogram, say $2 > x_1 > 0, |x_1 - x_2| \leq 1$, and the crease is the line $x_1 = x_2$. Then it is clear that, in the family parametrised by ϵ, the quantities $D(p)$ are not uniformly bounded, for p in an arbitrarily small neighbourhood of the origin. As we explained in Section 3, the essential difficulty in the proof of Theorem 3 is to show that we cannot have a family behaving in this way close to $(0,0)$ unless the $D(p)$ are also large

for some large p. The same discussion applies at the vertices. It is clear from our description of the asymptotics of the model solutions that for them the quantity $E(t)$ is unbounded as $t \to \infty$. Thus, in the setting of Section 5, we cannot obtain an *a priori* bound on $E_{\max}$ by "local" considerations around the vertex. The essential difficulty in the proof in Section 5 is to show that we cannot have a situation where $E(t)$ is large for some range of t but nevertheless $E(t)$ is bounded for very large t: in particular we cannot have a blow-up of the curvature, modelled on our explicit solutions, unless the $u^{(\alpha)}$ are already unbounded in the interior of P.

References

[1] Anderson, M.T. *Canonical metrics on 3-manifolds and 4-manifolds* Asian Jour. Math. **10** 2006 127-163

[2] Calderbank, D.M.J. and Pedersen, H. *Self-dual Einstein metrics with torus symmetry* Jour. Diff. Geom. **60** 485-521 2002

[3] Donaldson, S.K. *Scalar curvature and stability of toric varieties* Jour. Differential Geometry **62** 289-349 2002

[4] Donaldson, S.K. *Interior estimates for solutions of Abreu's equation* Collectanea Math. **56** 103-142 2005

[5] Donaldson, S.K. *Extremal metrics on toric surfaces: a continuity method* To appear in *Jour. Differential Geometry*

[6] Donaldson, S.K. *A generalised Joyce construction for a family of nonlinear partial differential equations* Preprint

[7] Joyce,D.D. *Explicit construction of self-dual 4-manifolds* Duke Math. J. **77** 519-552 1995

[8] Szekelyhidi, G. *Optimal test configurations for toric varieties* arxiv 07092687

[9] Zhou, B. and Zhu, X. *A note on K-stability of toric manifolds* arxiv 07060505